



Universidade de Brasília  
IE – Departamento de Estatística

**AVALIAÇÃO DO DESEMPENHO DAS UNIVERSIDADES BRASILEIRAS:  
UMA APLICAÇÃO DA ANÁLISE ENVOLTÓRIA DE DADOS**

Victor Hugo Lima de Almeida  
Hugo Batista Gomes

Orientador: José Angelo Belloni

Brasília  
2013



Universidade de Brasília  
IE – Departamento de Estatística  
Estágio supervisionado 2

Victor Hugo Lima de Almeida  
Hugo Batista Gomes

**AVALIAÇÃO DO DESEMPENHO DAS UNIVERSIDADES BRASILEIRAS:  
UMA APLICAÇÃO DA ANÁLISE ENVOLTÓRIA DE DADOS**

Orientador: José Angelo Belloni

Monografia de final de curso a ser submetida ao Departamento de Estatística da Universidade de Brasília como requisito parcial para a obtenção do título de Bacharel em Estatística.

Brasília

Dezembro de 2013

À Cida, Ane, Vinícius, Victoria e Vital, pelo  
apoio e carinho.

Victor Hugo Lima de Almeida

## **AGRADECIMENTOS**

Aos professores e colegas, pela convivência e aprendizado ao longo desse anos.

Ao professor Belloni, pelos ensinamentos que tornaram esse trabalho possível.

## SUMÁRIO

INTRODUÇÃO .....	1
1 CONSIDERAÇÕES INICIAIS .....	3
1.1 Censo da educação superior .....	3
1.1.1 O respaldo legal .....	3
1.1.2 Evolução da Metodologia e coleta .....	4
1.2 Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES ....	4
2 OBJETIVOS .....	6
2.1 Objetivo geral .....	6
2.2 Objetivos específicos .....	6
3 METODOLOGIA .....	7
3.1 Banco de dados .....	7
3.1.1 Obtenção das variáveis.....	8
3.2 Análise Exploratória .....	9
3.3 A matriz de correlações .....	16
3.4 Análise Multivariada de Dados.....	21
3.5 Análise em Componentes Principais – ACP.....	23
3.5.1 A Geometria do método ACP .....	24
3.5.2 O enfoque algébrico .....	26
3.6 Descobrimo as componentes principais.....	29
3.7 Análise Envoltória de Dados (DEA) .....	34
3.8 A seleção do modelo.....	42
3.8.1 As instituições eficientes do modelo <i>DEA3</i> .....	45
3.8.2 As Facetas obtidas e suas projeções.....	49
CONCLUSÕES E RECOMENDAÇÕES .....	55

REFERÊNCIAS.....	56
APÊNDICE A – INSTITUIÇÕES DA EDUCAÇÃO SUPERIOR.....	59
APÊNDICE B – A MATRIZ DE CORRELAÇÕES.....	61
APÊNDICE C - RELAÇÃO DAS INSTITUIÇÕES UTILIZADAS.....	66
APÊNDICE D – CORRELAÇÕES DAS VARIÁVEIS COM OS ÍNDICES DE EFICIÊNCIA ...	69
APÊNDICE E – VALORES OBSERVADOS E FOLGAS.....	70
APÊNDICE F – FRAÇÕES DE PRODUTOS POR INSUMO .....	73
APÊNDICE G – METAS PROPORCIONAIS E GLOBAIS .....	74

## RESUMO

Esta pesquisa consiste na avaliação do desempenho das universidades brasileiras sob o ponto de vista do critério da eficiência produtiva. A metodologia proposta baseia-se no uso iterativo de técnicas de Análise Exploratória de Dados e Análise Envoltória de Dados (DEA) e foi aplicada ao conjunto das universidades brasileiras a partir de dados secundários do Censo da Educação Superior e resultados das avaliações de cursos realizadas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

Uma análise estatística exploratória dos dados identificou os fatores educacionais mais relevantes na descrição das atividades universitárias e suas inter-relações. Características da Análise Envoltória de Dados viabilizaram o tratamento conjunto de variáveis descritoras desses fatores em uma análise global das instituições que, ao mesmo tempo, contempla as especificidades de cada universidade em respeito ao seu projeto institucional.

Foram avaliadas 100 universidades brasileiras selecionadas a partir de um critério de oferta de programas regulares de ensino de graduação e pós-graduação. Treze delas foram consideradas tecnicamente eficientes e compõem o conjunto de referência das demais. Para cada uma das demais a metodologia propôs um conjunto de ações e estratégias de melhoria da produtividade baseado em metas de produção obtidas a partir da projeção de cada instituição na fronteira de eficiência.

O acréscimo nas atividades de pós-graduação indica a principal alteração que deve ser efetuada nos projetos acadêmicos visando o crescimento da produtividade.

## INTRODUÇÃO

O universo das Instituições de Ensino Superior (IES) no Brasil é caracterizado por ser multifacetado. Existem diferentes organizações acadêmicas, sendo elas caracterizadas por diferentes projetos de ensino. Algumas instituições, como as faculdades, têm um maior enfoque na graduação, ofertando cursos tanto de bacharelado quanto de licenciatura, embora não atuem com pesquisa; outras, como os centros federais de educação tecnológica, apesar do enfoque similar na graduação, são responsáveis por realizar pesquisa. Devido a essa diferença de projetos, mensurar a eficiência de todas as IES brasileiras – de forma conjunta – sugere um método de padronização, que não é o enfoque adotado por este trabalho. Logo, o caminho que deve ser considerado é efetuar a distinção entre as organizações acadêmicas, ganhando em poder de comparabilidade. Tendo isso em mente, coube-nos realizar a escolha de uma organização acadêmica que englobasse o máximo que o ensino superior tem a oferecer, ou seja, a graduação (bacharelado, licenciatura ou formação tecnológica) e a pós-graduação *stricto sensu* (mestrado e doutorado) e *lato sensu* (*Master in business Administration* - MBAs e especialização), além da pesquisa e, conseqüentemente, publicações. Sendo assim, a escolha coube pela universidade. Mas qual delas apresenta um melhor desempenho? Qual é mais eficiente? Ou será que existem diferentes modelos de eficiência no território nacional? O seguinte trabalho intenta responder tais questionamentos.

Devemos ainda destacar que além da forma presencial de ensino, onde o aluno deve possuir frequência igual ou superior a 75%, ainda é possível o ingresso na modalidade ensino a distância (EAD). Conseqüentemente, mesmo dentro das universidades irão existir distintos projetos, uns voltados para a graduação com maior enfoque em cursos EAD, outros com uma sólida base na pós-graduação. Para medir a qualidade do ensino da graduação em meio a tais diferenças, o instituto nacional de pesquisas educacionais Anísio Teixeira (Inep) e o Ministério da Educação (MEC) criaram o Índice Geral de Cursos (IGC), calculado anualmente com base numa média dos conceitos de curso de graduação ponderada pelo



número de matriculados mais os conceitos da pós-graduação do ensino superior (também ponderado pelo número de matriculados).

Deste modo, o principal objetivo do trabalho é desenvolver uma metodologia de avaliação das instituições, particularmente das universidades, de uma maneira precisa e segura, fazendo uso da análise envoltória de dados. É desejável que tal metodologia seja eficiente em detectar padrões de associações entre as variáveis, além de ser capaz de distinguir os diferentes projetos acadêmicos desenvolvidos pelas universidades brasileiras. O objetivo ainda se desdobra em descobrir o quão factível é utilizar dados do ensino à distância para compor a análise, contemplando ainda a utilidade da metodologia de análise em componentes principais para a obtenção dos principais construtos compositores de uma instituição de ensino superior.

Para um maior grau de contextualização remetê-mos o leitor ao Apêndice A onde será possível obter a definição das organizações acadêmicas existentes no Brasil no ano de 2011.

## **1 CONSIDERAÇÕES INICIAIS**

### **1.1 Censo da educação superior**

Realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) - autarquia vinculada ao Ministério da Educação (MEC) - por meio da Diretoria de Estatísticas Educacionais (Deed), o censo da Educação superior é uma pesquisa de caráter declaratório e de participação obrigatória das instituições de ensino superior (IES). A coleta de dados é realizada através do acesso *online* ao sistema eletrônico de informações (Censup) e remete ao exercício acadêmico do ano anterior ao da coleta. As informações obtidas são tangentes a IES, cursos de graduação e sequenciais de formação específica além de alunos e docentes vinculados a tais cursos. A base de dados de referência a ser utilizada é o cadastro da instituição no Ministério da Educação (e-MEC).

O principal objetivo do censo é oferecer aos reitores, pró-reitores e dirigentes das IES, aos gestores das políticas educacionais e à sociedade em geral um “retrato da educação superior brasileira”.

Ao término da coleta e do período de verificação o Inep disponibiliza os microdados, o resumo técnico e a sinopse estatística objetivando garantir acessibilidade e uma maior transparência. O banco de dados utilizado neste trabalho é, em sua grande parte, obtido através desses microdados.

Desde o final da década de 1950 o censo já ocorria de maneira periódica e sistemática, sendo atualizado e aprimorado de maneira contínua desde então.

#### **1.1.1 O respaldo legal**

A constituição Brasileira de 1988 em seu artigo de número 21 dispôs sobre a necessidade da União de organizar e manter os serviços oficiais de estatística, dessa maneira diversos decretos e portarias versam sobre tal tema, delegando e outorgando as

responsabilidades nele contida. De acordo com a lei nº 9448 de 14 de março de 1997 é finalidade do Inep, entre outras, organizar e manter o sistema de informações e estatísticas educacionais, planejar, orientar e coordenar o desenvolvimento de sistemas e projetos de avaliação educacional.

A competência da Deed de, juntamente com as IES, planejar, promover e efetuar a coleta de dados emana do Decreto nº 6317, de 20 de dezembro de 2007, já a obrigatoriedade do preenchimento por parte das IES é reconhecido no Decreto nº 6425, de 4 de abril de 2008, antes era usada uma portaria para vincular a obrigação. O uso das informações contidas no cadastro do MEC (e-MEC) como a base de dados de referência a ser utilizada no censo é corroborada pela Portaria Normativa nº 40 de 12 de dezembro de 2007, republicada em 29 de dezembro de 2010.

Investidas de todo o arcabouço legal supracitado, as informações passadas pelas instituições, após o término do censo, passam a figurar como oficiais.

### **1.1.2 Evolução da Metodologia e coleta**

A partir do ano de 2009, o censo passou a coletar separadamente as informações sobre docentes e discentes visando diminuir eventuais erros no preenchimento e ganhar maior confiabilidade nas informações prestadas pelas IES, tornando os dados obtidos mais robustos. No censo de 2011 passou a ser obrigatório o preenchimento do cadastro de pessoa física (CPF) do discente, tentando com isso evitar a duplicidade de cadastros.

A coleta de dados ocorre de forma eletrônica através do endereço do Censup.

## **1.2 Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES**

Criada em 11 de julho de 1951 com o nome de Campanha Nacional de Aperfeiçoamento de pessoal de nível superior, a atual CAPES tinha como o objetivo “assegurar a existência de pessoal especializado em quantidade e qualidade suficientes para

atender às necessidades dos empreendimentos públicos e privados que visam ao desenvolvimento do país”. O País passava pelo período desenvolvimentista do segundo governo Vargas e era essencial a obtenção de pessoal especializado nas mais diversas áreas do conhecimento, de forma que durante tal período foram adotadas práticas, que seguem até os dias atuais, como o intercâmbio, a entrega de bolsas de estudo, contratação de professores estrangeiros e o estímulo à cooperação entre instituições de ensino superior. Com chegada da ditadura a coordenação passa a definir e regulamentar os cursos de pós-graduação nas universidades brasileiras, ganhando ainda o poder e meios para intervir na qualificação do corpo docente destas.

Após muitas reformas e reestruturações, no ano de 2007 a lei 11502 é aprovada e cria-se a nova CAPES que tem como funções coordenar o alto padrão do sistema nacional de pós-graduação brasileiro além de induzir e fomentar a formação inicial e continuada para professores da educação básica.

Atualmente a coordenação de aperfeiçoamento de pessoal de nível superior é uma fundação do ministério da educação (MEC) cujo papel na expansão do ensino de pós-graduação é de vital importância em todos os estados da federação. Assim as atividades da CAPES podem ser agrupadas em:

- Avaliação da pós-graduação *stricto sensu* (mestrado e doutorado);
- Acesso e divulgação da produção científica;
- Investimentos na formação de recursos de alto nível no país e no exterior;
- Promoção da cooperação científica internacional;
- Indução e fomento da educação inicial e continuada de professores da educação básica nos formatos presencial e a distância.

## **2 OBJETIVOS**

### **2.1 Objetivo geral**

O objetivo geral do trabalho é a avaliação das instituições de educação superior brasileiras, particularmente das universidades, fazendo uso da Análise Envoltória de Dados e levando em consideração a sua estrutura, informações acadêmicas do fluxo escolar do alunado e medidas de qualidade dos cursos e programas.

### **2.2 Objetivos específicos**

- Identificar os principais fatores educacionais e analisar suas inter-relações;
- Encontrar as universidades consideradas eficientes na transformação dos recursos em resultados acadêmicos;
- Mensurar a ineficiência das demais instituições através das suas projeções à fronteira de eficiência;
- Estabelecer metas de produção acadêmica para cada universidade, identificando ações e estratégias de melhoria da produtividade.

### 3 METODOLOGIA

#### 3.1 Banco de dados

O banco de dados é composto por informações das 190 universidades brasileiras ativas no ano de 2011, sendo obtido integralmente no ano de 2012 através do preenchimento do censo da educação superior. Todos os dados são autoinformados e cabe à instituição de ensino superior (IES) preenchê-lo de maneira fidedigna. Como o referido censo fornece apenas informações relativas ao quadro de graduação foi necessário recorrermos a outras bases, como por exemplo, a da coordenação e aperfeiçoamento de pessoal de nível superior (CAPES), com o intuito de obtermos variáveis relacionadas à pós-graduação e a pesquisa. Variáveis relacionadas à pesquisa não foram consideradas devido à dificuldade de obtê-las.

Assim, o banco utilizado contém variáveis relativas à:

- Alunos (número de matrículas, vagas, concluintes, ingressantes, entre outras);
- Titulação do corpo docente (número de doutores, mestres, professores com especialização, sem pós-graduação e sem graduação);
- Categoria administrativa (privada, pública federal, estadual e municipal);
- Estado e região;
- Biblioteca (quantidade de assentos, quantidade de comutação, quantidade de acervos e usuários capacitados);
- Financeiro (receita própria, despesa com docentes, despesa com pesquisa e investimento e outras);
- Quantidade de cursos ofertados;
- Indicadores da qualidade da graduação e pós-graduação (conceito médio graduação, conceito médio mestrados, conceito médio doutorado);
- Indicador da qualidade da universidade (índice geral de cursos - IGC);
- Número de matriculados e concluintes nos cursos de mestrados e doutorados.

Totalizando 39 variáveis.

Algumas observações quanto aos dados devem ser feitas. 1) A variável matrícula não é obtida de forma direta, é calculada como sendo a soma entre cursando e concluinte. 2) Grande

parte das variáveis oriundas de alunos e curso (na graduação) está dividida quanto à modalidade de ensino em presencial e ensino a distância (EAD) – que são projetos de ensino díspares. 3) Antes mesmo de efetuar análises estatísticas é possível notar que existem certos padrões entre algumas variáveis, algumas têm a característica de mensurar o tamanho das universidades, como as variáveis relacionadas aos discentes (matrícula, vagas, concluintes) e a quantidade de docentes; outras indicam a qualidade como o IGC e os conceitos médios.

### **3.1.1 Obtenção das variáveis**

A maioria das variáveis que compõe o banco de dados, por serem totais, não foi obtida de maneira direta, sendo necessário o seu cálculo. Assim, para obter o número total de matrículas, por exemplo, é necessário irmos até o banco de dados que contém todos os alunos do ensino superior no ano de 2011, solicitar aqueles que constam com o status de concluinte ou cursando e realizar a soma por curso, em seguida vincula-se o curso à respectiva universidade e realiza-se a soma de todas as matrículas em todos os cursos obtendo o total de matrículas naquela instituição. De maneira geral as variáveis referentes a aluno e docentes são obtidas de forma similar, enquanto as variáveis conceito (da graduação e da pós) e o IGC foram obtidos para cada universidade de forma direta. A obtenção das variáveis relacionadas aos discentes da pós-graduação foi obtida no site da CAPES, cabendo aqui uma observação: nem todas as universidades foram encontradas, com isso pode-se concluir que dentre as 190 universidades algumas não ofertam cursos de pós-graduação (ou não o fizeram no ano de 2011) de maneira que tais universidades não serão consideradas ao longo do estudo, tendo em vista que a principal característica de uma universidade, considerada neste trabalho, é a dualidade do ensino na graduação e na pós-graduação aliada à qualidade. Em tal banco ainda consta a cisão entre mestrado e mestrado profissional, que acreditamos que pode ser unida em uma única variável.

Um ponto importantíssimo nesse começo do estudo é a consideração do fato de que, na graduação, as universidades ofertam cursos tanto presenciais quanto à distância e que tal diferenciação, apesar de importante, introduz uma discrepância nos dados. A variável ingresso ilustra bem o caso. Segundo o censo superior o ingresso em uma IES se dá nas formas de ingresso por processo seletivo (dividido em vestibular, ENEM e outro tipo de seleção) e

outras formas de ingresso (ingresso por PECG e outro tipo de ingresso). Tais variáveis são binárias e informadas junto com o aluno de forma que não é possível o preenchimento em mais de um dos cinco campos acima apresentados. Para calcular o total de ingressantes é necessário somar tais variáveis por universidades. É possível, através da modalidade de ensino de cada curso, separar as duas variáveis iniciais (ingresso por processo seletivo e outras formas de ingresso) em presenciais e EAD. Decidimos considerar apenas por esse aspecto não importando a forma de ingresso. Como foi dito anteriormente, o aspecto EAD de tal variável pode ser problemático no sentido em que é uma variável que apresenta mediana igual a zero, ou seja, em metade das universidades no ano de 2011 não houve ingresso em cursos à distância.

### 3.2 Análise Exploratória

A separação das universidades por categoria administrativa auxilia no entendimento da dinâmica do ensino superior brasileiro, enquanto que a contagem de universidades por unidade federativa demonstra como ocorre sua distribuição no território nacional. Os dados estão resumidos na Tabela 1 e Tabela 2.

Tabela 1 – Número de universidades segundo a categoria administrativa

Categoria Administrativa		Número de Universidades	Total
Privada			88
Pública			102
	Federal	59	
	Estadual	37	
	Municipal	6	
Total			190

Fonte: Censo Superior da Educação (2011)

Através da análise de tais tabelas é factível ver o quão díspare é a distribuição das universidades no território nacional. Tome como exemplo o estado de São Paulo que, sozinho, possui mais universidades que as regiões Norte e Centro-Oeste somadas. Note ainda que 41,58% das universidades do Brasil estão concentradas na região sudeste. Parte desse arranjo se deve a própria pungência da região e principalmente do estado em que a universidade está inserida, portanto é de se esperar que São Paulo, por ser um dos principais centros urbanos brasileiros, abrigue uma atmosfera multifacetada, tanto econômica quanto social e tecnologicamente e que tal conjuntura demanda um maior número de universidades. Dessa



maneira é de esperar que as principais unidades federativas tenham um número maior de tais instituições. Além do mais, o envolvimento da instituição com a sociedade local e as disparidades regionais representam diferenças sociais, econômicas e culturais que não estão diretamente mensuradas no banco embora possam estar refletidas nas variáveis disponíveis.

Tabela 2 - Número de universidades segundo as unidades da federação

Região	Unidades da Federação	Número de Universidades	Total	% por Região
Norte	Acre	1	16	8,4
	Amazonas	3		
	Amapá	2		
	Pará	5		
	Rondônia	1		
	Roraima	2		
	Tocantins	2		
Nordeste	Alagoas	3	35	18,4
	Bahia	8		
	Ceará	6		
	Maranhão	2		
	Paraíba	3		
	Pernambuco	5		
	Piauí	2		
	Rio Grande do Norte	4		
	Sergipe	2		
Centro-Oeste	Distrito Federal	2	14	7,4
	Goiás	4		
	Mato Grosso	3		
	Mato Grosso do Sul	5		
Sudeste	Espirito Santo	1	79	41,6
	Minas Gerais	22		
	São Paulo	38		
	Rio de Janeiro	18		
Sul	Paraná	14	46	24,2
	Santa Catarina	13		
	Rio Grande do Sul	19		
Total			190	100

Fonte: Censo Superior da Educação (2011)

Analisando sob o prisma da categoria administrativa temos mais universidades públicas do que privadas no Brasil. Tal resultado não necessariamente informa sobre o quão democrático é o ensino superior brasileiro, pois uma pessoa de baixa renda pode vir a adentrar no ensino superior em instituições privadas devido a programas sociais, o que é um mérito das políticas públicas e não do fato de se ter mais ou menos universidades privadas.

Como a maioria das variáveis do estudo são divididas em presenciais e EAD é necessário entender o comportamento dentro destas duas realidades. Podemos então dividir as universidades em três grandes grupos.

O primeiro é composto por 73 instituições que ofertam vagas em cursos à distância e estes são preenchidos pelos ingressantes, temos aqui a Universidade de Brasília, Estácio de Sá e a Universidade Federal de Santa Catarina, por exemplo.

O segundo grupo é caracterizado pelas universidades que não têm ingressantes EAD já que não ofertam vagas para tal, é composto por 111 universidades e entre seus representantes temos Universidade de São Paulo, Federal do Mato Grosso e Federal de Ouro Preto.

O último grupo é formado por seis universidades que não ofertaram vagas para cursos à distância no ano de 2011, mas que tiveram ingressantes em tal modalidade. É preenchido pelas universidades federais de Viçosa, Mato Grosso do Sul, Rondônia e Amapá, além da Estadual do Ceará e a Universidade Castelo Branco. Tal grupo configura um elemento de estranheza nos dados tendo em vista que é impossível um aluno ingressar em uma vaga inexistente, fazendo-nos acreditar que tal evento foi um erro de preenchimento e não reflete a realidade, o que aponta – mais uma vez – na direção da retirada das variáveis EAD do banco.

Outro ponto importante é que a diferenciação quanto à modalidade de ensino cria *outliers* em variáveis importantes como vagas, matrículas e concluintes. Em outras palavras, as universidades que ofertam um grande número de cursos EAD tendem a se tornarem pontos muito acima dos demais, o que a tornaria uma unidade eficiente na leitura do DEA sem necessariamente o ser. O diagrama de caixa (*boxplot*) é a ferramenta gráfica ideal para a visualização de pontos discrepantes, do grau de variabilidade – observado pelo tamanho da “caixa” – além de fornecer medidas de amplitude e média. Na Figura 1 o *boxplot* foi feito levando em consideração as diferentes categorias administrativas para a variável TITULADO GRD, que nada mais é do que o total de concluintes em cursos da graduação. Pode-se perceber que a variabilidade dentro da universidade pública federal é semelhante a da universidade privada, todavia esta apresenta oito pontos discrepantes, ou seja, fora dos limites do *boxplot*, enquanto aquela não apresenta nenhum. Quais seriam esses pontos e qual a razão de estarem tão longe do intervalo citado?

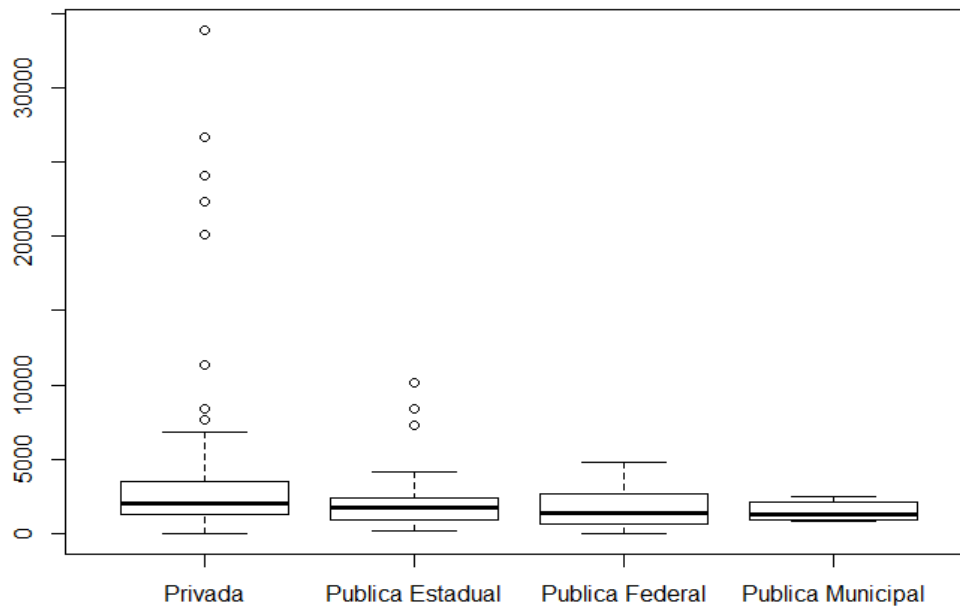


Figura 1 - *Boxplot* da variável TITULADO GRD por categoria administrativa

A resposta para a primeira pergunta está na Tabela 3, onde é possível visualizar as maiores universidades privadas em termos de concluintes na graduação. O comportamento da matrícula EAD em algumas universidades é claramente influenciador. A Universidade do Norte do Paraná, por exemplo, graduou 26.648 alunos no ano de 2011 dos quais 24.429 foram em cursos EAD, ou seja, 92% dos concluintes foram de cursos à distância.

Tabela 3 - Universidades privadas com maior número de titulados na graduação

Universidade	TITULADO GRD	Modalidade de ensino		EAD/TITULADO GRD
		Presencial	EAD	
Universidade Paulista	33.892	29.901	3.991	0,12
Universidade do Norte do Paraná	26.648	2.219	24.429	0,92
Universidade Estácio de Sá	24.113	23.384	729	0,03
Universidade Luterana do Brasil	22.328	5.483	16.845	0,75
Universidade Nove de Julho	20.135	19.338	797	0,04
Universidade Castelo Branco	11.331	812	10.519	0,93
Universidade Bandeirante de SP	8.411	8.411	0	0,00
Universidade Católica de MG	7.632	7.548	84	0,01

Fonte: Censo Superior da Educação (2011)

Um maior número de matriculados no ensino a distância pode ser visto na Universidade Luterana do Brasil e na Universidade Castelo Branco. O mesmo comportamento é repetido para as variáveis matrícula, inscritos, ingressantes de maneira que o a modalidade de ensino à distância se mostra uma “formadora de *outliers*”.

Concernente às universidades públicas municipais e estaduais podemos perceber que existe uma redução na variabilidade dos titulados perante o quadro anterior (universidades federais e privadas), sendo a variância mínima nas universidades municipais. Todavia, persistem alguns pontos discrepantes na categoria administrativa estadual, que são oriundos das universidades de São Paulo, do Tocantins e da estadual paulista Júlio de Mesquita Filho. Como podemos ver na Tabela 4, a universidade do Tocantins é composta majoritariamente por cursos à distância enquanto que nas demais universidades o que ocorre é o oposto.

Tabela 4 - Universidades estaduais com maiores matrículas (totais)

Universidade	TITULADO GRD	Modalidade de ensino	
		Presencial	EAD
Universidade de São Paulo (USP)	8.438	8.438	0
Universidade estadual paulista Júlio de Mesquita Filho	7.313	7.313	0
Universidade do Tocantins	10.109	0	10.109

Fonte: Censo Superior da Educação (2011)

Na Tabela 5 é possível verificar algumas estatísticas descritivas remetentes à variável TITULADO GRD e suas modalidades de ensino. Note que pelo menos 50% das universidades não possuem concluintes em cursos EAD (mediana é zero), tal fato somado a inserção de *outliers* nos leva a questionar se realmente é útil – para a nossa análise – a separação de variáveis pela modalidade de ensino.

Tabela 5 - Estatísticas descritivas – Número de titulados

Estatísticas descritivas	TITULADO GRD	Modalidade de ensino	
		Presencial	EAD
Mínimo	0,00	0,00	0,00
1° Quartil	932,75	821,50	0,00
Mediana	1.817,50	1.721,50	0,00
Média	2.766,23	2.264,95	501,28
3° Quartil	2.868,25	2.530,25	98,00
Máximo	33.892,00	29.901,00	24.429,00
Desvio padrão	4.225,96	3.228,37	2.421,95

Os seguintes gráficos de setores mostram que existe um grande número de vagas EAD sendo oferecidas – apesar de poucas universidades o fazerem – em contraponto ao baixo número de matrículas ocupadas por estas, ou seja, as universidades que ofertam vagas EAD o fazem muito acima da demanda o que reforça os questionamentos sobre a utilização da informação sobre educação superior à distância nesta pesquisa.

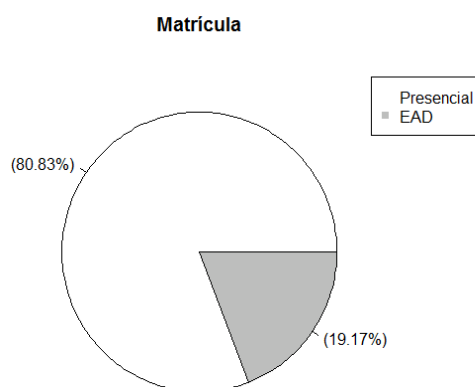


Figura 2 - variável matrícula (MATRICULA GRD)

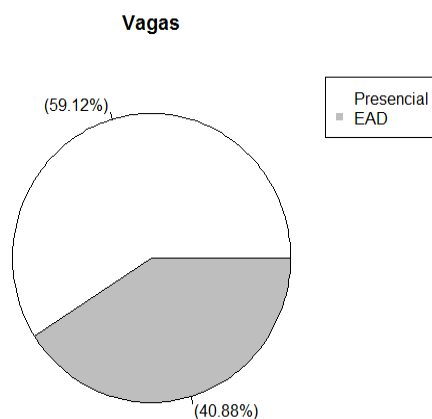


Figura 3 - variável vaga (VAGA TOTAL)

O próximo ponto a ser verificado é o comportamento do corpo docente em relação às categorias administrativas. As universidades federais são aquelas que, conforme o esperado,

apresentam maior quantidade de docentes doutores, seguida pelas estaduais e por fim pelas privadas. Um fator que chama a atenção é o fato de que justamente as federais apresentam em seu quadro um maior número de docentes sem pós-graduação. A quantidade de professores sem formação é ínfima e acreditamos poder ser somada com os docentes que não possuem pós-graduação sem perda de generalidade. Na Tabela 6 é possível visualizar o total do corpo docente dividido por titulação ao longo das categorias administrativas.

Tabela 6 - Titulação do corpo docente segundo a categoria administrativa

Titulação	Categoria Administrativa				Total
	Privada	Pública Estadual	Pública Federal	Pública Municipal	
Sem graduação	2	13	0	0	15
Sem pós-graduação	2.964	3.358	6.361	125	12.808
Especialização	20.935	9.035	4.145	1.195	35.310
Mestrado	32.127	13.214	20.943	1.681	67.965
Doutorado	17.489	21.622	47.275	749	87.135
Total	73.517	47.242	78.724	3.750	203.218

Fonte: Censo Superior da Educação (2011)

O comportamento de algumas universidades serve para ilustrar a disparidade existente no banco de dados. Tome por exemplo a UNIP. Essa universidade apresenta, na maioria das variáveis indicadoras do porte da instituição, valores muito superiores a qualquer outra instituição não tendo, portanto, padrão de comparação de mesmo porte no sistema. A caracterização da UNIP como *outlier* em algumas das variáveis pode sinalizar relações entre variáveis que inexistem. Por outro lado, essa instituição apresenta ensino de pós-graduação de pequeno porte e uma proporção de docentes com doutorado inferior a 20% indicando um projeto pedagógico concentrado no ensino de graduação.

As universidades também se distinguem pelos projetos acadêmicos adotados. O projeto acadêmico de uma instituição está refletido nas variáveis presentes no banco de dados e, em particular, nas relações entre elas. Razões entre variáveis do ensino de graduação e de pós-graduação revelaram que estas atividades têm importâncias relativas muito diferentes de uma universidade para a outra. A titulação do professor é um fator diferenciador das instituições que, indiretamente, também reflete as particularidades dos projetos acadêmicos.

Como seria então a relação entre os docentes e a modalidade de ensino? Como as variáveis de conceito se relacionariam com as universidades que fornecem, principalmente, cursos a distancia? Para respondermos todas essas perguntas é necessário a análise da matriz de correlações.

### 3.3 A matriz de correlações

Para efetuar a análise de uma matriz de correlações é necessário observar o agrupamento das variáveis, como esses agrupamentos se relacionam uns com os outros e como as variáveis relacionam-se dentro deles. No Apêndice B, é possível conferir as principais correlações que compõe a matriz de covariância obtidas das análises aqui apresentadas.

O primeiro grupo identificado é relacionado à graduação, contendo variáveis como matrículas, vagas, concluintes, ingressantes, falecidos, quantidade de cursos entre outras; o segundo grupo é aquele que remete à titulação dos docentes; o terceiro é relativo ao conceito da instituição – tanto na graduação como na pós-graduação; o quarto grupo é justamente o que contém variáveis referentes à pós-graduação (matrículas e concluintes) e o último grupo é o que contém informações relacionadas a publicações, existindo ainda uma variável isolada que é a quantidade de acervo nas bibliotecas.

Percebe-se que dentro do primeiro grupo existe, de maneira geral, uma forte correlação entre as variáveis, além de que, algumas outras possuem baixa correlação com as demais como, por exemplo, a variável alunos falecidos. É importante ressaltar que existem algumas variáveis que, apesar de um alto grau de correlação, não necessariamente serão incluídas no estudo, pois apresentam caráter pouco explicativo como matrículas trancadas, transferidos dentro da mesma universidade e desvinculados.

No segundo grupo as correlações apresentam-se de maneira menos intensa, e até negativa em alguns pontos, o que já era de se esperar. O docente sem graduação relaciona-se de forma negativa com os demais (com exceção dos docentes doutores), enquanto que o docente sem pós se relaciona de forma positiva, porém com pouca intensidade. A maior correlação encontrada é a de docentes mestres com docentes especializados (0,756).

O grupo dos conceitos é caracterizado por baixa correlação entre suas variáveis componentes, assim podemos inferir que o fato de uma universidade possuir uma excelente graduação não tem implicância na qualidade da pós-graduação, dessa maneira pode-se distinguir dois projetos educacionais distintos nas universidades brasileiras: o primeiro voltado mais para o ensino da graduação e o segundo refere-se a uma ênfase no ensino da pós. Entre os conceitos, a maior correlação é a do mestrado com o doutorado, corroborando o que

foi dito acima. Nota-se também que o índice geral de cursos (IGC) se relaciona fortemente (0,94) com o conceito médio da graduação, o que já era esperado, tendo em vista que o cálculo do IGC utiliza tal conceito ponderado pelo número de alunos da graduação.

Se, por um lado, as correlações entre variáveis da graduação e da pós-graduação são relativamente fracas, por outro lado, observaram-se associações extremamente fortes entre as variáveis do mestrado e suas correspondentes para o doutorado, expressas por coeficientes de correlação superiores a 0,9. Além disso, as variáveis do mestrado e do doutorado apresentam um padrão de associação semelhante em relação às demais variáveis.

Agora que já conhecemos o funcionamento das correlações intragrupos é necessário observarmos o comportamento intergrupos.

O grupo relacionado à graduação se relaciona, de maneira generalizada, fracamente com os demais, sendo que a única correlação forte que temos é com o professor mestre. Por sua vez o grupo da titulação dos docentes apresenta correlação forte entre o professor mestre e o número de cursos da graduação e o professor doutor e as despesas das universidades. O grupo dos conceitos é caracterizado por ter baixa correlação com os demais grupos. Observa-se, no ensino da pós-graduação, forte correlação com o professor doutor (acima de 0,86) e com a quantidade de acervos da biblioteca.

A variável representativa da biblioteca, número de títulos de livros (acervo), tem sua correlação mais forte com o número de docentes com doutorado (0,80), correlações fortes com as variáveis da pós-graduação (maiores que 0,63) e mais fracas com as variáveis da graduação (menores que 0,43). Além disto, associa-se positivamente com a proporção de professores com doutorado e negativamente com as demais titulações do docente. Neste sentido, essas correlações parecem indicar que a existência de programas de pós-graduação, com a consequente maior titulação do corpo docente, é um fator indutor de uma biblioteca com maior acervo.

Deste ponto pode-se concluir que o professor mestre é aquele que está intimamente ligado ao ensino da graduação, o que é explicado pela sua forte correlação com o alunado e com o número de cursos das universidades, enquanto que o professor doutor é mais atuante no ensino da pós-graduação sendo, teoricamente, mais dispendioso mantê-lo, outrossim voltamos novamente para ideia dos diferentes projetos educacionais. Veja bem, se a universidade não tem um ensino sólido da pós-graduação, não vale a pena – financeiramente falando – manter



um corpo de docentes doutores, já que estes estão voltados para o projeto de formação de doutores e mestres.

Concomitantemente aos resultados obtidos cabem duas perguntas, a primeira é: sob a luz da análise da matriz de correlações, é necessário usarmos todas as variáveis? A resposta é não. Ora, percebemos que existem diversas variáveis que não agregam informação alguma para a nossa análise, seja por se relacionarem de maneira baixíssima com as variáveis mais importantes do estudo, seja pela sua natureza. Variáveis relativas a biblioteca (com exceção para quantidade de acervos) se encaixam no primeiro caso, enquanto variáveis financeiras no segundo. Para esta última cabe um adendo: não a consideraremos devido a sua essência “enganadora” no sentido que os dados não são fidedignos, muitas vezes os números ali indicados são apenas dos repasses do governo e não do valor real. Considerando-a estaríamos acrescentando um fator de estranheza nos dados nos levando a tomar conclusões errôneas.

O segundo questionamento que deve ser respondido remete ao problema citado anteriormente das variáveis EAD: É necessário utiliza-las? Depreende-se da análise que os níveis de correlação para variáveis importantes são sempre menores se as considerarmos sob o aspecto EAD. Adicione a isso os demais resultados das seções anteriores apontando na direção de sua retirada. Por conseguinte, opta-se por preservar o conjunto de dados não fazendo uso das variáveis que remetem à modalidade de ensino à distância.

Ao término deste capítulo ainda chegamos à conclusão de, como dito anteriormente, não considerarmos universidades que não ofereçam o ensino da pós-graduação, ganhando, deste modo, um maior nível de comparabilidade com as universidades restantes, e diminuindo o número de universidades utilizadas para 101. Entretanto, ainda existe uma universidade que, por não apresentar indicadores de qualidade, deve ser retirada da análise para não obtermos resultados discrepantes. Desse modo, opta-se por retirar a Universidade de São Paulo (USP) e trabalhar apenas com as 100 universidades restantes. No Apêndice C é possível obter o rol de todas as 100 universidades utilizadas neste trabalho, além de sua categoria administrativa.

Definidas as instituições que participaram da análise, selecionou-se previamente um conjunto de 55 variáveis. Todas as variáveis descritoras das atividades de extensão e algumas variáveis relativas à pesquisa científica, mesmo constituindo informações relevantes, foram abandonadas ao longo desse trabalho por insuficiência de dados. Da mesma forma as variáveis financeiras relativas a custos e receitas das instituições foram desconsideradas por

apresentarem inconsistências e não serem objeto de auditoria pública que validassem as informações.

Com essas restrições resultou um conjunto de 25 variáveis representativas do fluxo do alunado na graduação (6 variáveis) e pós-graduação (4), do corpo docente (4), da biblioteca (1), indicadores de qualidade das atividades (4) e do endereçamento da instituição (6), conforme detalhado no Quadro 1. Essa seleção considerou a necessidade de representar os principais aspectos envolvidos na atividade universitária.

A análise de correlações lineares permitiu uma primeira leitura das relações existentes, todavia não possibilitou o alcance dos objetivos desta análise de dados. Algumas perguntas permanecem sem respostas adequadas:

- Como identificar as associações entre os conjuntos de variáveis que descrevem os principais fatores educacionais presentes no banco de dados?
- Quais as possibilidades de agregação de variáveis?

Visando responder a estes questionamentos, identificar as estruturas de relações entre as variáveis e encontrar os principais fatores determinantes das diferenças entre as instituições faremos uso de uma ferramenta da análise multivariada: a Análise em componentes principais.

Quadro 1 - Apresentação das variáveis e sua nomeação no banco de dados

<p>Ensino de Graduação:</p> <p>Número de cursos de graduação – QT CURSO  Número de vagas oferecidas – VAGA TOTAL  Número de inscritos no vestibular – INSC VEST  Número de alunos matriculados – MATRICULA GRD  Número de diplomados – TITULADO GRD  Número de ingressantes – ING TOTAL</p>
<p>Ensino de Pós-Graduação - Mestrado:</p> <p>Número de alunos matriculados – MATRICULA MESTRADO  Número de diplomados – TITULADO MESTRADO</p>
<p>Ensino de Pós-Graduação - Doutorado:</p> <p>Número de alunos matriculados – MATRICULA DOUTORADO  Número de diplomados – TITULADO DOUTORADO</p>
<p>Corpo Docente:</p> <p>Número total de professores – DOCENTE TOTAL  Número de professores com doutorado – DOC DOUTORADO  Número de professores com mestrado – DOC MESTRADO  Número de professores com graduação/especialização – DOC OUTROS</p>
<p>Biblioteca</p> <p>Número de títulos de livros – QT ACERVO</p>
<p>Indicadores de qualidade</p> <p>Conceito médio graduação – CONCEITO GRD  Conceito médio mestrado – CONCEITO MESTRADO  Conceito médio doutorado – CONCEITO DOUTORADO  Índice Geral de Cursos da IES (contínuo) – IGC</p>
<p>Endereçamento da IES</p> <p>Código da IES – CO_IES  Nome da IES – NO_IES  Categoria administrativa  Sigla  Unidade da Federação – UF  Região geográfica – REGIAO</p>

### 3.4 Análise Multivariada de Dados

O arcabouço de metodologias referente a um conjunto de dados composto de mensurações simultâneas de diversas variáveis é chamado de Análise multivariada sendo também conhecida como Análise multivariada de dados, ou análise de dados multivariados. Seu estudo consiste na aplicação de métodos e técnicas a um número grande de variáveis observadas simultaneamente em um conjunto de dados, concentrando-se no estudo das relações entre variáveis e grupos de variáveis. Existem diversos métodos multivariados, sendo alguns baseados em modelos probabilísticos, como aqueles que supõem a distribuição normal multivariada, enquanto outros são justificados apenas pelo uso dos dados, contudo, independentemente de sua origem, todos são implementados de maneira computacional. Entre suas diversas aplicações podemos citar:

- Redução no conjunto de dados/ simplificação estrutural: A representação do fenômeno ocorre da maneira mais simples possível, não havendo perda de informações preciosas. Objetiva tornar a interpretação mais fácil;
- Agrupamento: agrupar variáveis similares de acordo com suas características;
- Investigação da relação entre variáveis:
- Predição;
- Construção e teste de hipótese.

Existem diversas classificações para técnicas multivariadas. Algumas distinguem entre técnicas designadas para estudo de relações interdependentes e técnicas designadas para relações de dependência – aquelas utilizadas para prever uma variável em função das demais. Outra classifica as técnicas em relação ao número de populações e o número de conjunto de variáveis que está sendo estudado. A escolha das técnicas a serem utilizadas depende dos objetivos do estudo em questão e da natureza das variáveis envolvidas. No caso deste estudo estamos interessados em estudar as associações entre todas as variáveis, objetivando descrever o comportamento e as estruturas do banco de dados para, por fim, efetuar a simplificação da informação. Dessa maneira, devemos utilizar técnicas de relações de interdependência como, por exemplo, análise fatorial, análise de correspondência, análise de conglomerados e análise em componentes principais. Qual seria a técnica mais adequada para o nosso caso? Como dito acima a resposta pode ser obtida do cerne das variáveis do estudo. Nossas variáveis são todas quantitativas, logo as técnicas se reduzem a análise de

conglomerados, análise fatorial e análise de componentes principais. E dessas, qual seria a mais indicada?

A análise de conglomerados (ou análise de *cluster*) é um ferramental multivariado utilizado na combinação de variáveis tentando a formação de grupos que sejam homogêneos e compactos dentro de si, ou seja, as características internas são semelhantes para as variáveis que o compõe, enquanto que os grupos são heterogêneos entre si (características diferem de grupo para grupo). Como o nosso objetivo não é a criação de grupos e sim a seleção de variáveis, de forma que haja a menor perda possível de informação, tal procedimento não é adequado.

A análise fatorial e a análise em componentes principais (ACP) apresentam uma metodologia bastante similar, mas é de suma importância definir seus pontos divergentes. A análise fatorial apresenta um caráter plural no sentido em que ela pode ser feita de uma maneira exploratória ou confirmatória. Na análise fatorial exploratória a estrutura dos modelos fatoriais ou a teoria envolvida na construção de tais modelos não é conhecida *a priori*, assim os dados são utilizados para revelar/identificar tais estruturas. Por outro lado, na análise fatorial confirmatória a estrutura do modelo já é conhecida, sendo baseada em alguma teoria pré-existente. Dessa maneira não é factível utilizarmos análise fatorial de caráter confirmatório, tendo em vista que não estamos supondo nenhum comportamento ou modelo a ser seguido pelos dados. A análise fatorial exploratória e a análise em componentes principais visam reduzir a dimensionalidade dos dados, assim muitos autores colocam esta como um tipo daquela. Todavia existem algumas diferenças materiais entre os dois métodos: 1) Não é necessário ser feita nenhuma suposição na análise em componentes principais; 2) A ênfase em ACP é explicar a variância total enquanto que na análise fatorial é explicar as covariâncias; Por último podemos citar a discrepância na gênese destas análises. Enquanto que na ACP as componentes principais são combinações lineares das variáveis originais, na análise fatorial, as variáveis são expressas como uma combinação linear dos fatores.

De acordo com tudo que foi dito até aqui, conclui-se que o procedimento mais adequado para reduzir a dimensão dos dados, perdendo o mínimo de informação possível e não fazendo nenhuma suposição sobre eles é a análise em componentes principais.

### 3.5 Análise em Componentes Principais – ACP

A Análise em componentes principais, uma das mais conhecidas técnicas da análise multivariada, planeja simplificar um problema reduzindo-lhe o número de variáveis originais, de maneira que ao invés de se trabalhar com sua integralidade, trabalha-se com um reduzido número de novas variáveis, agora independentes, que são combinações lineares das variáveis originais (as chamadas componentes principais). Desta maneira a ACP esta concentrada em explicar as estruturas de variâncias/covariâncias dos dados através destas combinações.

As variáveis originais constituem um espaço de dimensão  $p$ , responsável por explicar toda a variância dos dados, todavia, grande parte desta variabilidade pode ser explicada por um número “pequeno” de  $r$  componentes principais. Assim, é possível substituímos as  $n$  mensurações das  $p$  variáveis por  $n$  mensurações das  $r$  componentes principais, não havendo perda de informação, visando sempre à maximização da variância das combinações lineares.

Podemos definir as componentes principais da seguinte maneira: A primeira componente principal é a combinação linear que apresenta variância máxima, a segunda componente é a combinação linear de máxima variância, ortogonal a primeira e assim sucessivamente. Uma ilustração bastante do uso da ACP é dada a seguir.

Considere um banco de dados onde há correlação entre as variáveis. Quando se fizer uma regressão linear, percebe-se multicolinearidade, o que acaba por aumentar o desvio padrão e proporcionar um modelo que não explique tanto a situação desejada. Uma solução seria utilizar componentes principais criando novas variáveis que não possuam correlação entre si, utilizando-as numa regressão realmente eficiente.

O trabalho inicial de ACP deve-se a Karl Pearson que propôs, em 1901, a projeção de um espaço  $p$  dimensional em um espaço linear ou bidimensional. Através da utilização de regressão linear propôs encontrar combinações lineares das variáveis originais que melhor se ajustassem aos dados. Com o advento dos computadores, Hotelling (1933) desenvolveu uma técnica de extração das componentes principais (originando-se aqui tal termo) tornando o método bastante popular. Devido a tal popularização, tradicionalmente, é atribuída a Hotelling a criação da técnica.

### 3.5.1 A Geometria do método ACP

Considere a situação em que temos diversas observações de modo que estas são correlacionadas. Quando se “plota” os desvios da média referentes a elas em um plano ( $X_1$ ,  $X_2$ ), é possível, através de uma rotação de  $\theta$  graus do eixo  $X_1$ , criarmos o eixo  $X_1^*$  (Figura 4) de modo que as projeções dos pontos  $x_1$  e  $x_2$  em tal eixo são dadas por:

$$x_1^* = \cos\theta x_1 + \sin\theta x_2$$

Assim,  $x_1^*$ , que é uma combinação linear das demais variáveis, é uma nova variável que apresenta variância maior do que a das variáveis  $X_1$  e  $X_2$ . Tal variância pode ser maximizada variando-se o ângulo  $\theta$ , obtendo tal valor computacionalmente.

Note que apesar de  $x_1^*$  explicar uma maior porção da variância do conjunto de dados ele não a explica por completo, doravante, é possível identificar um segundo eixo, correspondente a uma segunda variável, que terá a tarefa de explicar a máxima variabilidade não explicada por  $X_1^*$ . Tal eixo –  $X_2^*$  – será ortogonal ao primeiro, sendo assim, se o ângulo entre  $X_1^*$  e  $X_1$  é  $\theta$ , este é o mesmo para  $X_2^*$  e  $X_2$  o que faz com que as duas novas variáveis sejam independentes (não correlacionadas). A combinação linear que dará origem a variável  $x_2^*$  é:

$$x_2^* = \cos\theta x_2 - \sin\theta x_1$$

Chamamos de componentes principais os novos eixos gerados ( $X_1^*$  e  $X_2^*$ ) e de valores de componentes os valores que caracterizam esses eixos ( $x_1^*$  e  $x_2^*$ ). Observe que através de tal método a variável  $X_1^*$  explica, unicamente, grande parte da variância dos dados, e, que efetuando a soma de sua variância com a variância de  $X_2^*$  obtemos a variância total. Tal resultado é importantíssimo no que tange ao uso da ACP como uma técnica de diminuição dimensional dos dados, pois agora é possível trabalhar apenas com  $X_1^*$ .

É imprescindível que o mesmo raciocínio seja ampliado para um espaço dimensional maior, não há limitação nesse campo; então, havendo a necessidade, em vez de se trabalhar com um número grande de  $p$  variáveis, pode-se trabalhar com um número menor (preferencialmente muito menor) de apenas  $m$  variáveis que seriam a combinação linear das variáveis originais.

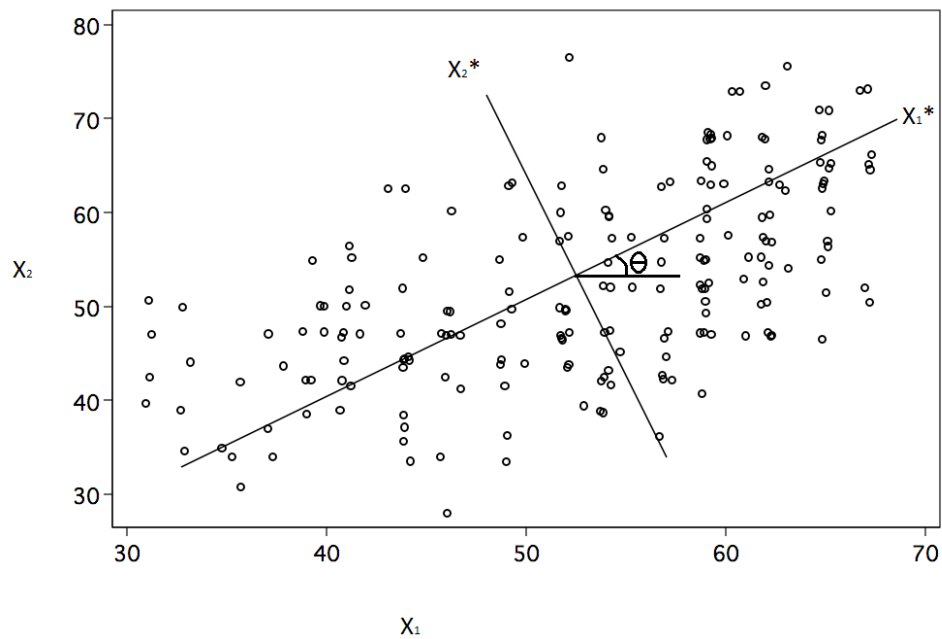


Figura 4 - Determinação das componentes principais através da rotação dos eixos

Como dito anteriormente, a variância explicada pela primeira componente principal é a maior possível, a segunda é também a maior possível que não é explicada pela primeira variável, a terceira é a maior possível que não é explicada pelas duas primeiras e assim por diante, de modo que a explicação da variância estará em ordem decrescente as componentes principais. A quantidade de informação perdida por uma Análise em componentes principais, quando há a tentativa de se reduzir o espaço dimensional dos dados originais (utilizando-se apenas a primeira componente, por exemplo), é explicada pela quantidade de variância descartada quando há a seleção – nesse caso a perda seria a soma das variâncias explicadas pelas demais componentes.

Note que para se realizar tal método é necessário se fazer ou uma correção da média para zero ou uma padronização dos dados. Isso dependerá dos dados originais, se há a necessidade de se impor igual importância as variáveis então o segundo método deve ser utilizado, isso porque se a variância de uma variável é substancialmente maior do que as demais sua importância na formulação dos eixos será maior, se houver a padronização então a mesma importância será dada. Nesse caso, uma maior correlação nos dados proporcionaria uma melhor possibilidade de redução, quanto menos correlacionados, menor é a explicação dada pelos eixos, e conseqüentemente, maior será o número de componentes principais.



### 3.5.2 O enfoque algébrico

Estamos interessados em realizar uma transformação linear de maneira que seja possível reproduzir a variabilidade contida nas variáveis originais. Logo, é necessário definir o que é variabilidade em uma conjuntura multivariada. Quando se tem uma única variável sua variabilidade é expressa em termos de sua variância (quando existe). No caso de um vetor, a variância é dada por uma matriz: a matriz de covariâncias. Uma maneira de reduzir tal matriz a um único número é calculando a variância total, que é a soma das variâncias das variáveis originais, que no caso da matriz de covariâncias, seria a soma dos elementos de sua diagonal, ou seja, o seu traço.

Seja  $\mathbf{X} = (x_1, x_2, \dots, x_p)$  o vetor das  $p$  variáveis dos dados originais. A matriz de covariâncias é dada pela  $E(\mathbf{X}\mathbf{X}') = \mathbf{V}$  e, como foi dito acima, a variância total é calculada fazendo-se  $\sum_{i=1}^p \text{var}(x_i) = \text{traço}(\mathbf{V})$ . Como o principal objetivo da ACP é encontrar uma combinação linear das variáveis originais, devemos considerar  $\boldsymbol{\gamma}' = (\gamma_1, \gamma_2, \dots, \gamma_p)$  o vetor de pesos da combinação linear dos dados originais, e conseqüentemente  $\mathbf{Y} = \boldsymbol{\gamma}'\mathbf{X}$  designa o vetor das componentes principais, cuja variância é dada por  $E(\mathbf{Y}\mathbf{Y}') = E(\boldsymbol{\gamma}'\mathbf{X}\mathbf{X}'\boldsymbol{\gamma})$  ou  $\boldsymbol{\gamma}'\mathbf{V}\boldsymbol{\gamma}$ . Por conseguinte obtêm-se as equações:

$$y_1 = \gamma_{11} x_1 + \gamma_{12} x_2 + \dots + \gamma_{1p} x_p$$

$$y_2 = \gamma_{21} x_1 + \gamma_{22} x_2 + \dots + \gamma_{2p} x_p$$

...

$$y_p = \gamma_{p1} x_1 + \gamma_{p2} x_2 + \dots + \gamma_{pp} x_p$$

Os pesos são estimados de forma que obedeçam às restrições:

1. A primeira componente principal responde pela variância máxima dos dados, a segunda componente principal responde pela variância máxima dos dados não explicada pela primeira componente, e assim sucessivamente; Note que se as  $p$  variáveis originais forem correlacionadas, então o número de componentes principais será menor do que  $p$ , o que proporciona uma redução nas variáveis sem perda de informação;
2.  $\gamma_{i1}^2 + \gamma_{i2}^2 + \dots + \gamma_{ip}^2 = 1 \quad i = 1, 2, \dots, p$  ou de forma vetorial  $\boldsymbol{\gamma}'\boldsymbol{\gamma} = 1$ ;

$$3. \gamma_{i1}\gamma_{j1} + \gamma_{i2}\gamma_{j2} + \dots + \gamma_{ip}\gamma_{jp} = 0, \forall i \neq j.$$

A segunda condição é usada para fixar a escala das novas variáveis, porquanto é possível aumentar a variância de uma combinação linear mudando a escala dos pesos. Tome por exemplo a primeira componente principal, podemos aumentar sua variância em quatro vezes assumindo que  $\gamma_{11} = 2\gamma_{11}, \gamma_{12} = 2\gamma_{12}$  e assim sucessivamente.

A terceira condição impõe a ortogonalidade entre os novos eixos, ou seja, que as componentes principais sejam independentes umas das outras.

O problema da maximização pode ser resolvido através dos multiplicadores de Lagrange:

$$Z = \gamma'V\gamma - \lambda(\gamma'\gamma - 1);$$

Sendo  $\lambda$  o multiplicador de Lagrange, faz-se então a derivação parcial em relação ao vetor de pesos, obtendo-se a seguinte equação:

$$\frac{\partial Z}{\partial \gamma} = 2V\gamma - 2\lambda\gamma$$

Quando igualada a zero resulta na solução final, que é:

$$(V - \lambda I)\gamma = 0 \rightarrow |V - \lambda I| = 0$$

Ou seja, uma equação polinomial em  $\lambda$  de grau  $p$ , possuindo  $p$  raízes. Agora, considere que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$  sejam as  $p$  soluções para a equação. Tais valores nada mais são do que os autovalores da matriz  $V$ . Obtemos assim até  $p$  componentes principais, cada uma para cada autovalor não nulo. Assim, o primeiro autovetor  $\gamma_1$ , correspondente ao primeiro autovalor  $\lambda_1$  é obtido resolvendo-se:

$$(V - \lambda_1 I)\gamma_1 = 0 \text{ e } \gamma_1'\gamma_1 = 1$$

Multiplicando-se a primeira parte por  $\gamma_1'$  temos:

$$\gamma_1'(V - \lambda_1 I)\gamma_1 = 0 \rightarrow \gamma_1'V\gamma_1 = \lambda_1\gamma_1'\gamma_1 = \lambda_1$$

Visto que temos a restrição  $\gamma_1'\gamma_1 = 1$ . Desta maneira a variância da nova variável  $y_1$  é igual ao autovalor  $\lambda_1$ . De forma similar pode ser mostrado que os vetores de peso  $\gamma_2', \gamma_3', \dots, \gamma_p'$  são os autovetores correspondentes aos autovalores  $\lambda_2, \lambda_3, \dots, \lambda_p$  da matriz de covariância  $V$ . Logo, o problema de obter os pesos para as combinações lineares que darão

origem às componentes principais é reduzido a encontrar os autovetores (que designarão os pesos) e os autovalores (que serão as variâncias das componentes principais).

Depois de feita a análise de componentes principais, a seleção de variáveis a serem utilizadas pode ser feita de diferenciadas maneiras, dependendo principalmente de três atributos: A natureza do estudo, seus objetivos e o grau de precisão desejado. Dentre as diversas formas de seleção as mais utilizadas são:

1. Coletar variáveis suficientes até se obter 80% da variância dos dados;
2. Reter as variáveis com autovalores maiores que a média dos autovalores, ou seja,  $\sum_{i=1}^p \frac{\lambda_i}{p}$ ;
3. Através da análise do gráfico *scree plot*, onde é possível se procurar a quebra entre os valores altos e baixos dos autovalores; procura-se por um “cotovelo”.

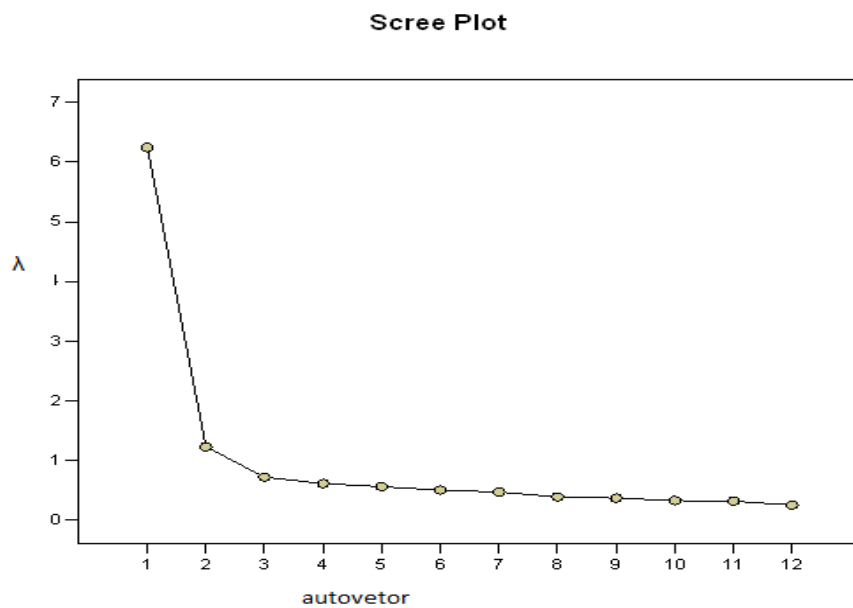


Figura 5 - Exemplo de *scree plot*

Como dito acima há outros métodos para fazer a seleção das variáveis, porém são mais complexos sendo necessários apenas para um critério mais rígido. Testes de hipótese podem ser aplicados com o intuito de se selecionar os autovalores maiores, a análise paralela de Horn (1965) também é uma alternativa que pode ser utilizada.

### 3.6 Descobrimo as componentes principais

Ao buscar fatores independentes que, em conjunto, reproduzam a variabilidade existente nos dados originais, a ACP identificou os principais fatores que explicam as diferenças entre as universidades.

Das 55 variáveis utilizadas nesta análise exploratória, 25 são resultantes do processo de seleção. As variáveis agregadas e suas parcelas não podem ser usadas simultaneamente em ACP porque introduzem dependência linear entre elas. As análises fatoriais foram realizadas usando as variáveis em sua forma mais desagregada, introduzindo as agregações como variáveis suplementares. O conjunto reduziu-se a 16 variáveis sobre as quais se realizou uma análise em componentes principais cujos resultados estão resumidos na Tabela 7 e Tabela 8. Destacaram-se cinco componentes principais que, em conjunto, explicam 93,03% da variabilidade total existente nos dados e correspondem a fatores educacionais distintos.

Tabela 7 - Explicação da variabilidade pelas componentes principais

Componentes Principais	Autovalores	Explicação da variabilidade	
		% Variância	% Variância Acumulada
1	6,809	42,56	42,56
2	5,577	34,86	77,41
3	1,135	7,09	84,50
4	0,917	5,73	90,24
5	0,447	2,79	93,03
6	0,346	2,17	95,20
7	0,195	1,22	96,42
8	0,166	1,04	97,45
9	0,154	0,96	98,41
10	0,093	0,58	99,00
11	0,061	0,38	99,38
12	0,047	0,29	99,67
13	0,033	0,21	99,88
14	0,009	0,06	99,94
15	0,005	0,03	99,97
16	0,004	0,03	100

O primeiro desses dois grandes fatores, associado à primeira componente principal, explica 42,56% da variabilidade total e representa uma informação comum a quase todas as variáveis, um elemento de estabilidade entre elas, que expressa o construto mais importante para definir as diferenças entre as instituições. Essa primeira componente apresenta associação linear positiva forte com todas as variáveis, a exceção dos indicadores da

qualidade que são variáveis que não dependem do tamanho da instituição (vide Tabela 8), refletindo, a princípio, o tamanho, a história e as relações de cada instituição com a sociedade, construto que foi denominado porte da instituição. A Tabela 8 apresenta as correlações lineares entre as variáveis originais e as componentes principais, auxiliando na compreensão do sentido destas. Conclui-se, então, que o construto mais importante para explicar as diferenças entre as instituições, sob o ponto de vista das variáveis disponíveis, é o porte da instituição e está associado a 1ª componente principal.

Tabela 8 – Correlação entre as variáveis originais e as componentes principais.

Variáveis	Componentes Principais						
	1	2	3	4	5	6	7
MATRICULA PRESC GRD	0,84	-0,47	0,06	0,00	0,00	-0,20	-0,10
VAGAS PRESC GRD	0,73	-0,62	0,05	0,10	0,05	-0,03	0,13
ING PRESC GRD	0,76	-0,56	0,06	0,09	0,06	-0,19	-0,02
TITULADO PRESC GRD	0,80	-0,50	0,09	-0,07	-0,04	-0,25	-0,03
QT CURSOS	0,82	-0,42	-0,08	-0,04	-0,09	0,11	-0,10
DOCENTE OUTROS	0,71	-0,58	-0,02	0,02	0,17	0,28	0,15
DOCENTE MESTRADO	0,75	-0,55	-0,06	0,07	0,06	0,24	-0,04
DOCENTE DOUTORADO	0,73	0,59	-0,18	0,05	0,11	0,15	-0,07
CONCEITO GRADUACAO	0,11	0,36	-0,12	0,91	-0,09	-0,03	-0,01
CONCEITO MESTRADO	0,38	0,57	0,66	0,06	0,05	-0,05	0,24
CONCEITO DOUTORADO	0,33	0,57	0,68	0,02	0,05	0,13	-0,25
MATRICULA MESTRADO	0,55	0,75	-0,27	-0,11	0,12	-0,07	0,00
TITULADO MESTRADO	0,57	0,76	-0,24	-0,10	0,12	-0,07	0,00
MATRICULA DOUTORADO	0,58	0,79	-0,14	-0,07	0,04	-0,04	0,03
TITULADO DOUTORADO	0,57	0,79	-0,08	-0,05	0,02	-0,02	0,04
QT ACERVO	0,73	0,30	-0,01	-0,13	-0,59	0,09	0,06

A segunda componente principal, responsável por 34% da variabilidade total, contrapõe dois conjuntos de variáveis. Ela tem associação negativa com as variáveis da graduação e com professores sem doutorado e positiva com as variáveis da pós-graduação, com os indicadores da qualidade da pós-graduação e com professores com doutorado. Essa componente reforça a primeira grande divisão das variáveis em dois grupos já percebida na análise de correlações lineares, separando as atividades de graduação das de pós-graduação, como pode ser visto na Figura 6, reafirmando a necessidade de se trabalhar com o corpo docente desagregado segundo a titulação. Tal figura apresenta o primeiro plano fatorial expresso a partir das duas primeiras componentes principais; cada ponto deste gráfico expressa a correlação linear entre estas e a variável original representada. Desta forma a

distância de cada ponto à origem corresponde ao coeficiente de determinação múltiplo entre a variável original e as duas primeiras componentes.

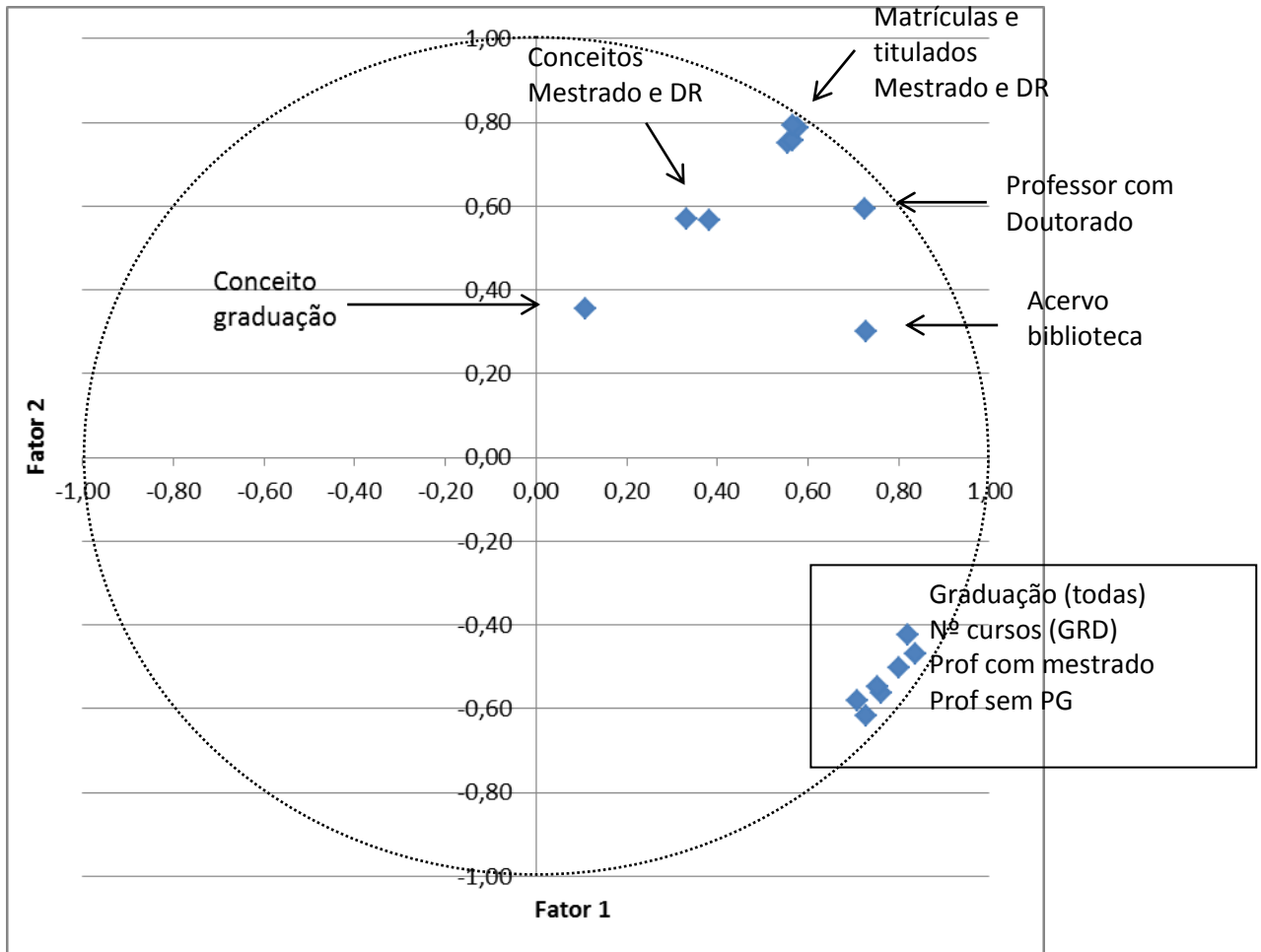


Figura 6 - Separação das variáveis nas duas primeiras componentes principais

A divisão obtida pelo primeiro plano fatorial também pode ser observada para as universidades. Analisando a Figura 7 é possível observarmos três conjuntos distintos de instituições. O primeiro é aquele guiado pelas grandes instituições públicas brasileiras, capitaneado pelas universidades estadual de São Paulo, de Minas Gerais, do Rio Grande do Sul (o uso da USP é meramente ilustrativo), apresentando projetos acadêmicos voltados para o ensino da pós-graduação, de maneira que quanto mais longe do centro da figura mais consolidado é esse projeto. O segundo é composto pelas grandes universidades privadas, instituições com um grande porte e um ensino forte apenas na graduação, destacamos a UNIP, Estácio de Sá e a Nove de Julho. O último grupo, referente as universidades que estão no centro, geralmente são de pequeno porte e a informação disponível não permitiu estabelecer diferenças significativas nos projetos acadêmicos.

Deste modo foi inferido que o segundo construto importante para diferenciar as instituições expressa diferenças de projeto acadêmico no que diz respeito à ênfase atribuída às atividades de graduação em relação às atividades de pós-graduação.

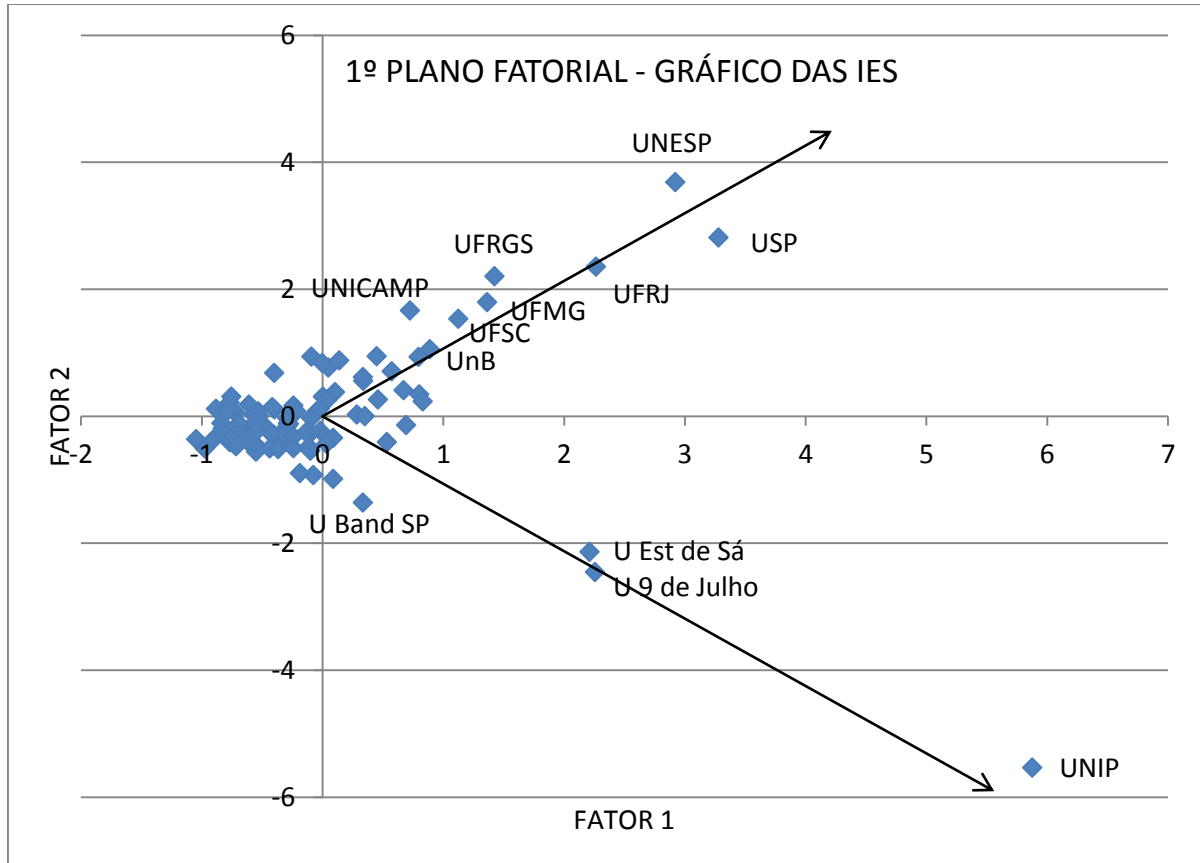


Figura 7 - As universidades no primeiro plano fatorial

A terceira e a quarta componentes principais estão associadas aos indicadores da qualidade das atividades de graduação e pós-graduação e explicam conjuntamente 12,82% da variabilidade total. A terceira componente (7,09% da variabilidade) é dominada pelos indicadores da qualidade da pós-graduação enquanto a quarta é dominada pelo indicador de qualidade da graduação. A necessidade de utilizar duas componentes principais para expressar os indicadores da qualidade aponta uma diferença clara entre eles (indicadores da graduação e da pós-graduação) já expressa nas diferenças entre as atividades que eles avaliam.

Os coeficientes de determinação expressam a qualidade da representação de cada variável pelo conjunto das componentes e podem ser encontrados na Tabela 9. Podemos dizer, por exemplo, que as duas primeiras componentes explicam 95% da variável quantidade de doutores titulados, enquanto que a primeira explica apenas 32,2%. Tal informação é de suma importância uma vez que podemos ter um alto grau de explicação pelas componentes

principais e um baixo grau de representação nas variáveis principais o que é um indicativo de que uma nova componente deve ser considerada.

Tabela 9 - Qualidade da representação das variáveis.

Variáveis	Componentes principais						
	1	2	3	4	5	6	7
MATRICULA PRESC GRD	0,701	0,92	0,924	0,924	0,924	0,965	0,975
VAGAS PRESC GRD	0,529	0,91	0,912	0,923	0,925	0,926	0,942
ING PRESC GRD	0,580	0,90	0,901	0,909	0,913	0,951	0,951
TITULADO PRESC GRD	0,642	0,89	0,901	0,906	0,908	0,970	0,971
QT CURSOS	0,673	0,85	0,859	0,860	0,868	0,879	0,890
DOCENTE OUTROS	0,503	0,84	0,839	0,840	0,869	0,946	0,969
DOCENTE MESTRADO	0,569	0,87	0,874	0,879	0,882	0,938	0,940
DOCENTE DOUTORADO	0,527	0,88	0,910	0,912	0,925	0,946	0,951
CONCEITO GRADUACAO	0,012	0,139	0,154	0,988	0,996	0,997	0,997
CONCEITO MESTRADO	0,147	0,466	0,904	0,907	0,910	0,912	0,972
CONCEITO DOUTORADO	0,111	0,435	0,901	0,901	0,904	0,920	0,982
MATRICULA MESTRADO	0,307	0,868	0,943	0,955	0,969	0,974	0,974
TITULADO MESTRADO	0,322	0,893	0,950	0,960	0,974	0,980	0,980
MATRICULA DOUTORADO	0,334	0,955	0,973	0,978	0,980	0,981	0,982
TITULADO DOUTORADO	0,322	0,950	0,956	0,959	0,960	0,960	0,961
QT ACERVO	0,529	0,619	0,619	0,636	0,978	0,986	0,990

Observe que a variável acervo da biblioteca só estará bem representada considerando a quinta componente principal (2,79% da variabilidade total) e está parcialmente nas quatro componentes anteriores. Assim esta variável parece estar presente em todas as dimensões expressas nestas cinco primeiras componentes caracterizando-se como um elemento de infraestrutura que permeia todas as atividades acadêmicas representadas no banco de dados.

Consequentemente, podemos concluir, que a análise em componentes principais é uma ferramenta de extrema valia no que se refere a descobrir como funciona a composição das universidades brasileiras explicitando a interação entre variáveis e diferenciando os projetos de ensino de cada instituição.



### 3.7 Análise Envoltória de Dados (DEA)

A técnica DEA, que significa Análise Envoltória de Dados (do inglês, *Data Envelopment Analysis*) teve início com a tese de doutorado de Rhodes em sua obtenção do Ph.D supervisionada por Cooper e publicada em 1978 (RHODES, 1978). É uma técnica usada para estimar as eficiências de unidades organizacionais homogêneas, que usam um mesmo conjunto de recursos para produzir um mesmo conjunto de resultados, através de processos tecnológicos similares. Em uma análise *ex-post*, DEA considera valores observados de recursos utilizados e resultados alcançados para construir um espaço de possibilidades de produção, delimitados por uma fronteira linear por partes definida a partir dos melhores desempenhos. Unidades que se posicionam sobre essa fronteira são ditas eficientes, enquanto a ineficiência das demais é mensurada a partir de uma medida da distância à fronteira. A projeção de cada unidade ineficiente na fronteira de eficiência determina metas que caracterizam ações capazes de aumentar sua produtividade.

O método utiliza de um conjunto de entradas (*inputs*) e saídas (*outputs*) para medir a eficiência relativa de empresas, firmas ou qualquer outra unidade de serviço em questão, aqui tratadas por DMU's (do inglês, *Decision Making Units*). Através da identificação das unidades eficientes é possível identificar e estudar as unidades ineficientes para se descobrir formas de reduzir tais ineficiências.

Trata-se de um procedimento multivariado não paramétrico, que fornece dados quantitativos para possíveis direções com o intuito de otimizar o status de uma observação tida como não eficiente, comparando entradas e saídas sem suposições estatísticas.

Segundo Farrel (1957) uma DMU é dita ótima quando consegue eficiência técnica, que é ter *output* máximo para o seu conjunto de *inputs*, e eficiência alo cativa, que seria usar proporções ótimas, dado seu setor econômico e tecnológico.

As vantagens do DEA são observadas quando se leva em consideração que a análise é feita individualmente em cada DMU, e sua eficiência é encontrada baseada nos resultados das outras observações, leva-se em consideração também o fato de unidades de proporções menores não serem prejudicadas na análise, dependendo do modelo escolhido.

O estabelecimento de metas é bem proveitoso no foco empresarial, pois dá alternativas tangíveis de melhoramento, além de unidades com perfis diferentes poderem ser consideradas ótimas. O modelo também apresenta simples interpretação, e se aplicado ao longo do tempo é possível ver a evolução de uma unidade e interpretar quais os fatores realmente contribuem para um aperfeiçoamento da mesma.

Para conceituar formalmente a eficiência mensurada pela técnica DEA, considere cada organização sob avaliação como um sistema de produção múltipla, que transforma  $N$  itens de recursos, representados por um vetor de quantidades  $x = (x_1, x_2, \dots, x_N) \in \mathbf{R}_+^N$ , em  $M$  itens de resultados cujas quantidades estão representadas em um vetor  $y = (y_1, y_2, \dots, y_M) \in \mathbf{R}_+^M$ , determinando um plano de produção descrito pelo vetor  $(x, y) \in \mathbf{R}_+^{N+M}$ . Suponha que foram observados  $K$  planos de produção  $(x^k, y^k), k = 1, 2, \dots, K$  realizados por  $K$  organizações similares.  $x_{ki}$  é a quantidade do recurso  $i$  ( $i = 1, 2, \dots, N$ ) utilizada e  $y_{kj}$  é a quantidade do resultado  $j$  ( $j = 1, 2, \dots, M$ ) produzida pela  $DMU^k$ . A organização sob avaliação será simbolizada por  $DMU^0$ , representada nos modelos DEA pelo plano de produção  $(x^0, y^0)$ .

A Análise por Envoltória de Dados constrói, para cada  $DMU^0$ , uma medida de desempenho que tem a seguinte expressão:

$$E_0 = \frac{\sum_{j=1}^M p_j y_{0j}}{\sum_{i=1}^N q_i x_{0i}},$$

Onde  $p = (p_1, p_2, \dots, p_M)$  e  $q = (q_1, q_2, \dots, q_N)$  são os vetores de pesos utilizados para a agregação dos recursos e dos resultados da  $DMU^0$ , respectivamente. O desempenho resultante,  $E_0$ , é uma medida de produtividade da  $DMU^0$ , por configurar-se uma razão entre a produção agregada e o consumo agregado.

A utilização de um conjunto de pesos para cada DMU resulta em uma medida de desempenho específica para cada organização sob avaliação, que busca reproduzir os valores expressos no plano de produção realizado.

A determinação dos pesos para cada  $DMU^0$  é feita maximizando-se o valor da produtividade  $E_0$ , sujeito à restrição de que a produtividade de nenhuma DMU, calculada com os pesos da  $DMU^0$ , possa exceder um valor constante pré-fixado (usualmente tomado igual a 1). O seguinte problema de programação matemática pode ser usado para calcular tais pesos:

$$E_0 = \max \frac{\sum_{j=1}^M q_j y_{oj}}{\sum_{i=1}^N p_i x_{oi}}$$

$$s/a \quad E_k = \frac{\sum_{j=1}^M q_j y_{kj}}{\sum_{i=1}^N p_i x_{ki}} \leq 1 \quad \forall k = 1, 2, \dots, K \quad (1)$$

$$p_i \geq 0, \quad \forall i = 1, 2, \dots, N; \quad q_j \geq 0, \quad \forall j = 1, 2, \dots, M$$

Como a produtividade máxima observada será sempre igual a 1, a medida  $E_0$ , produtividade da  $DMU^0$ , pode ser dividida por essa produtividade máxima, constituindo-se, assim numa medida da eficiência relativa da  $DMU^0$ .

Na solução deste problema de otimização, a eficiência da  $DMU^0$  é maximizada sob a condição que a eficiência de cada uma das unidades não excede o valor 1. Assim, uma  $DMU^K$  será considerada eficiente, sob o ponto de vista da  $DMU^0$ , quando sua medida de eficiência relativa  $E_K$  (calculada com os pesos da  $DMU^0$ ) for igual a 1, e ineficiente quando esta medida for menor que 1.

O modelo definido em (1) caracteriza-se por um problema de programação fracional não convexo que pode ser reduzido a dois problemas de programação linear. Essa transformação se faz alterando a função objetivo em (1) e resulta em dois problemas de programação linear; um que mantém constante o agregado de recursos e busca maximizar o agregado de produção (2); e, outro, que minimiza o agregado de recursos, mantendo constante o agregado de produção (3). Esses problemas estão explicitados na Figura 8.

Os problemas (2) e (3) são chamados problemas dos multiplicadores por expressarem as taxas de substituição entre recursos e entre resultados (os pesos) que definem a faceta da fronteira de eficiência na qual é projetada a  $DMU^0$ . O conceito de desempenho que está sendo mensurado em cada um deles torna-se mais claro quando se analisam problemas equivalentes a (2) e (3). Os problemas (4) e (5) da Figura 9 correspondem aos problemas duais dos problemas (2) e (3), respectivamente.

DEA orientado para o consumo (problema dos multiplicadores)	DEA orientado para a produção (problema dos multiplicadores)
$\max_{p_i, q_j} \sum_{j=1}^M y_{0j} q_j$ <p>s/a</p> $\sum_{i=1}^N x_{0i} p_i = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq 0, \quad \forall i = 1, 2, \dots, N$ $q_j \geq 0 \quad \forall j = 1, 2, \dots, M$ <p style="text-align: center;">( 2 )</p>	$\min_{p_i, q_j} \sum_{i=1}^N x_{0i} p_i$ <p>s/a</p> $\sum_{j=1}^M y_{0j} q_j = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq 0, \quad \forall i = 1, 2, \dots, N$ $q_j \geq 0 \quad \forall j = 1, 2, \dots, M$ <p style="text-align: center;">( 3 )</p>

Figura 8 - Problema dos multiplicadores

DEA orientado para o consumo (problema do envelopamento)	DEA orientado para a produção (problema do envelopamento)
$\min \theta$ <p>s/a</p> $\sum_{k=1}^K z_k y_{kj} \geq y_{0j}, \quad \forall j = 1, 2, \dots, M$ $x_{0i} \theta - \sum_{k=1}^K z_k x_{ki} \geq 0 \quad \forall i = 1, 2, \dots, N$ $\theta \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$ <p style="text-align: center;">( 4 )</p>	$\max \lambda$ <p>s/a</p> $y_{0j} \lambda - \sum_{k=1}^K z_k y_{kj} \leq 0, \quad \forall j = 1, 2, \dots, M$ $\sum_{k=1}^K z_k x_{ki} \leq x_{0i}, \quad \forall i = 1, 2, \dots, N$ $\lambda \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$ <p style="text-align: center;">( 5 )</p>

Figura 9 - Problema do envelopamento

As regiões de viabilidade dos problemas (4) e (5) caracterizam, respectivamente, o conjunto de necessidades de consumo associado ao vetor de produção da  $DMU^0$  e o conjunto de possibilidades de produção associado ao seu vetor de consumo. Os escalares  $z_k$  são os coeficientes dos planos de produção  $(x_k, y_k)$  nas combinações lineares que definem a tecnologia de produção.

O valor ótimo para  $\theta$  no problema (4), que notaremos  $\theta^*$ , representa a contração equiproporcional máxima possível no vetor de recursos da  $DMU^0$ , mantendo-se constante o vetor de resultados observados. Se nenhuma contração equiproporcional for possível, então  $\theta^* = 1$  e a  $DMU^0$  é dita eficiente no conjunto de necessidades de consumo. Se  $\theta^* < 1$ , então a  $DMU^0$  é ineficiente e  $\theta^*$  é a medida de sua ineficiência, já que os recursos utilizados poderiam ser reduzidos equiproporcionalmente de  $x^0$  para  $\theta^* x^0$  sem redução nos resultados produzidos.  $\theta^*$  caracteriza-se como uma medida da eficiência orientada para a redução do consumo de recursos.

De maneira equivalente,  $\lambda^*$ , o valor ótimo do problema (5), é a expansão máxima possível no vetor de resultados da  $DMU^0$ , mantendo-se constante o vetor de recursos utilizados. Se  $\lambda^* = 1$ , então nenhuma expansão é possível e a  $DMU^0$  é dita eficiente no conjunto de possibilidades de produção. Se  $\lambda^* > 1$ , então a  $DMU^0$  é dita ineficiente e seus resultados podem ser expandidos de  $y^0$  para  $\lambda^* y^0$  sem acréscimo de recursos. O valor  $1/\lambda^*$  define a medida de eficiência da  $DMU^0$  orientada para o aumento da produção de resultados.

As medidas definidas nos problemas (4) e (5) se caracterizam por uma projeção radial (equiproporcional) sobre a fronteira, mantendo, portanto as proporções entre recursos e entre resultados observadas na  $DMU^0$ .

Os problemas de programação linear, apresentados nas Figura 8 e Figura 9 correspondem ao modelo chamado de CCR (Charnes, Cooper e Rhodes, 1978), também conhecido como CRS (do inglês, *Constant Returns to Scale*) e estabelece uma fronteira de retornos constantes a escala de produção. A hipótese de retornos constantes pressupõe tanto a possibilidade de expansão proporcional do plano de operação, quanto de contração proporcional desse plano até a origem.

A Figura 10 ilustra as medidas radiais de eficiência, salientando as orientações para a maximização da produção e para a minimização do consumo e as projeções radiais (equiproporcionais) para as fronteiras.

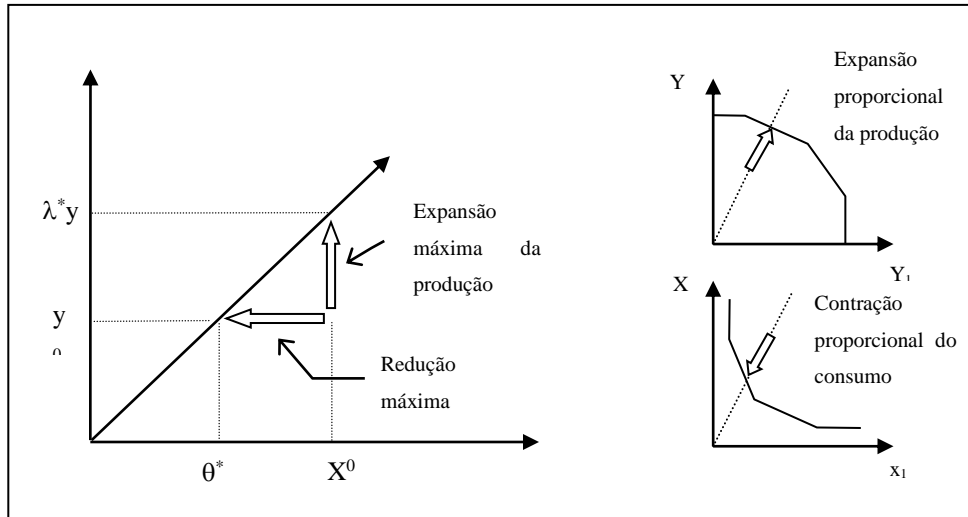


Figura 10 - Medidas de eficiência: orientações

O segundo modelo DEA, chamado BCC (Banker, Charnes e Cooper, 1984), também conhecido como VRS (do inglês, *Variable Returns to Scale*), utiliza uma fronteira de retornos variáveis, onde cada DMU ineficiente é projetada nas barreiras que mais atendem aos objetivos. A hipótese de retornos variáveis admite apenas a combinação convexa dos planos de operação observados, impedindo, assim, outras expansões e contrações desses planos.

As diferenças de fronteiras oriundas dos modelos CCR e BCC podem ser visualizadas na Figura 11. Os modelos se diferenciam pela introdução, no modelo BCC, da restrição de convexidade no problema do envelopamento que corresponde à introdução de uma nova variável no problema dos multiplicadores que expressa o afastamento da hipótese de retornos constantes. As Figura 12 e Figura 13 apresentam os problemas de programação linear associados ao modelo BCC.

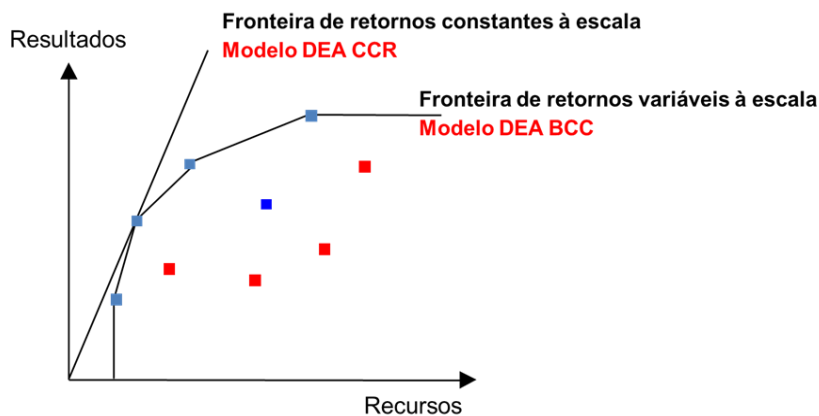


Figura 11 - Diferenças de fronteiras entre os dois modelos

Modelo BCC orientado para o consumo	
Retornos variáveis à escala - descarte forte de recursos e resultados	
$\max_{p_i, q_j} \sum_{j=1}^M y_{0j} q_j + v_0$ <p>s/a</p> $\sum_{i=1}^N x_{0i} p_i = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j + v_0 \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq \varepsilon, \quad \forall i = 1, 2, \dots, N$ $q_j \geq \varepsilon \quad \forall j = 1, 2, \dots, M$ $\varepsilon > 0,$ <p>(problema dos multiplicadores)</p>	$\text{mix} \theta + \varepsilon (\sum_{j=1}^M s_j^+ + \sum_{i=1}^N s_i^-)$ <p>s/a</p> $x_{0i} \theta - \sum_{k=1}^K z_k x_{ki} - s_i^- = 0, \quad \forall i = 1, 2, \dots, N$ $\sum_{k=1}^K z_k y_{kj} - s_j^+ = y_{0j}, \quad \forall j = 1, 2, \dots, M$ $\sum_{k=1}^K z_k = 1$ $\theta \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$ $s_j^+ \geq 0, \quad \forall j = 1, 2, \dots, M; \quad s_i^- \geq 0, \quad \forall i = 1, 2, \dots, N$ <p>(problema do envelopamento)</p>

Figura 12 - Modelo BCC orientado para o consumo

Conclui-se que para alcançar os objetivos estipulados nesse trabalho a melhor opção de modelo a ser utilizada é o BCC, pois ele respeita o tamanho das instituições, pré-requisito de muita valia para a nossa análise comparativa. Consequentemente a projeção radial do mesmo irá preservar os projetos acadêmicos, mantendo a proporcionalidade, além do que é importante ressaltar que os retornos não devem ser constantes, pois não é possível para uma universidade crescer desenfreadamente.

Modelo BCC orientado para a produção	
Retornos variáveis à escala - descarte forte de recursos e resultados	
$\min_{p_i, q_j} \sum_{i=1}^N x_{0i} p_i + v_0$ <p>s/a</p> $\sum_{j=1}^M y_{0j} q_j = 1$ $\sum_{i=1}^N x_{ki} p_i - \sum_{j=1}^M y_{kj} q_j + v_0 \geq 0, \quad \forall k = 1, 2, \dots, K$ $p_i \geq \varepsilon, \quad \forall i = 1, 2, \dots, N$ $q_j \geq \varepsilon \quad \forall j = 1, 2, \dots, M$ <p style="text-align: right;"><math>\varepsilon &gt; 0,</math></p> <p>(problema dos multiplicadores)</p>	$\max \lambda + \varepsilon \left( \sum_{j=1}^M s_j^+ + \sum_{i=1}^N s_i^- \right)$ <p>s/a</p> $y_{0j} \lambda - \sum_{k=1}^K z_k y_{kj} + s_j^+ = 0, \quad \forall j = 1, 2, \dots, M$ $\sum_{k=1}^K z_k x_{ki} + s_i^- = x_{0i}, \quad \forall i = 1, 2, \dots, N$ $\sum_{k=1}^K z_k = 1$ <p style="text-align: right;">Retornos variáveis</p> $\lambda \in R; \quad z_k \geq 0, \quad \forall k = 1, 2, \dots, K$ $s_j^+ \geq 0, \quad \forall j = 1, 2, \dots, M \quad s_i^- \geq 0, \quad \forall i = 1, 2, \dots, N$ <p>(problema do envelopamento)</p>

Figura 13 - Modelo BCC orientado para a produção

Relativamente à orientação do modelo, escolhemos orientar para a produção, tendo em vista que o objetivo do nosso estudo é maximizar os resultados sem alterar os recursos.



### 3.8 A seleção do modelo

A escolha do modelo DEA a ser utilizado baseia-se no procedimento de seleção de variáveis de Norman e Stoker (Norman e Stoker, 1991). Conhecida a classificação das variáveis em recursos ou resultados, o procedimento identifica aquelas que mais influenciam o desempenho de uma universidade e constrói um indicador único da eficiência produtiva que considera essas variáveis conjuntamente. Esse procedimento identifica os recursos e os resultados mais relevantes à mensuração do desempenho adotado através da construção de uma sequência de funções de desempenho.

O modelo DEA da primeira etapa de tal procedimento corresponde a um único resultado e um único recurso. Em cada uma das etapas seguintes são observadas as associações entre cada variável e o indicador da eficiência da etapa anterior para decidir se existem novas variáveis a considerar na função de desempenho. Essas novas variáveis podem representar um novo recurso ou um novo resultado, bem como decomposições de variáveis já presentes no modelo.

Se alguma variável que influencia a eficiência estiver ausente da função de desempenho, então o indicador da eficiência calculado deverá estar viesado com relação ao fator (insumo ou produto) descrito por esta variável. Deve-se esperar, nesse caso, associação estatística significativa entre o indicador da eficiência e essa variável, que é indicativa da possibilidade de entrada dessa variável na função de desempenho.

A relevância da decomposição de variáveis que estão presentes de forma agregada na função de desempenho é verificada através da análise das correlações entre as suas parcelas e o indicador da eficiência. Associações inversas entre parcelas e a eficiência são sinalizadoras da necessidade de decomposição da variável agregada.

O Apêndice D apresenta as correlações entre as variáveis originais e os modelos aqui estudados. Seguindo essa metodologia de seleção o primeiro modelo DEA a ser utilizado é aquele que apresenta apenas um insumo e um produto, ambos na forma mais agregada possível.

Executando o procedimento de Norman e Stoker e consoante com a ACP que identificou o tamanho das instituições como o principal fator de explicação da variabilidade,

optou-se por considerar como insumo inicial o total de docentes (DOCENTE TOTAL), e como produto gerado o total de titulados (TITULADO TOTAL), seja na graduação ou na pós-graduação. Consequentemente o primeiro modelo DEA obtido é da forma:

$$DEA_1 = \frac{TITULADO\ TOTAL}{DOCENTE\ TOTAL}$$

Cujo inverso é definido como a medida de eficiência

$$EFIC_1 = \frac{1}{DEA_1}$$

Calculando as correlações da variável  $EFIC_1$  com as demais é notado que existe uma associação inversa (correlações negativas) com as variáveis relativas ao fluxo da pós-graduação, aos conceitos e ao professor com doutorado, e relação direta (correlação positiva) com as variáveis do fluxo da graduação e os demais professores (que sabemos que atuam no ensino da graduação). Tais medidas nos informam que quanto mais docentes doutores, por exemplo, menos eficiente seria a universidade, o que não faz o mínimo sentido. Portanto, a medida  $EFIC_1$  não é adequada para mensurar o desempenho de uma universidade.

As correlações entre  $EFIC_1$  e as demais variáveis indicam dois caminhos possíveis a serem trilhados: o primeiro seria selecionar novas variáveis que considerem as diferenças entre o ensino da pós-graduação e da graduação, ou seja, que considerem os diferentes projetos educacionais das instituições, acarretando na desagregação do total de titulação. O segundo caminho sugerido é o acréscimo de uma variável que considere a qualidade do ensino oferecido nas instituições.

A análise de componentes principais considera que após o porte da universidade (nossa primeira componente principal) o principal fator de explicação da variabilidade é a diferença entre os projetos institucionais, assim, a ênfase dada ao ensino da graduação ou da pós-graduação deve ser considerada. Em seguida, a terceira e a quarta componente apontam na direção dos indicadores de qualidade, sendo estes um fator de explicação que deve ser considerado após a separação graduação/pós-graduação.

Para a construção do segundo modelo optou-se por seguir a ordem que a ACP traz, fazendo a desagregação da variável Total de titulados em duas componentes Titulados na graduação (TITULADO GRD) e Titulados na pós-graduação (TITULADO PG) dando origem ao modelo  $DEA_2$ :

$$DEA_2 = \frac{q_1 TITULADO GRD + q_2 TITULADO PG}{DOCENTE TOTAL}$$

Com sua respectiva medida de eficiência

$$EFIC_2 = \frac{1}{DEA_2}$$

A análise das correlações entre  $EFIC_2$  e as demais variáveis aponta para uma associação positiva tanto entre as variáveis de fluxo na graduação quanto na pós-graduação, tendo, deste modo, captado o fator de diferença que é explicado pela segunda componente principal. Todavia, tal modelo ainda não percebe os efeitos dos indicadores de qualidade, sejam eles da graduação ou da pós e conseqüentemente do próprio IGC, não conseguindo expressar a qualidade das atividades expressas nestes conceitos. As associações obtidas fornecem correlações baixíssimas (e em alguns casos negativas) entre  $EFIC_2$  e tais variáveis, indicando a necessidade de acrescentar indicadores de qualidade ao nosso modelo. Optou-se, então, pela entrada do IGC como um novo produto, justamente pelo fato de tal indicador ser uma síntese dos demais. O terceiro modelo obtido é então:

$$DEA_3 = \frac{q_1 TITULADO GRD + q_2 TITULADO PG + q_3 IGC}{DOCENTE TOTAL}$$

$$EFIC_3 = \frac{1}{DEA_3}$$

Pela análise das correlações de  $EFIC_3$  pode-se perceber que esse modelo tem um ajuste muito bom aos dados e não oferece nenhuma alternativa de mudança – qualquer acréscimo ou desagregação torna-o pior – além do mais, contempla o conceito de uma universidade eficiente explorado desde o início do trabalho fazendo com que este seja o modelo escolhido.

Com a escolha do modelo já efetuada, cabem dois adendos relativos a ela. Caso tivéssemos optado por não seguir a direção apontada pela ACP e partíssemos do modelo  $DEA_1$  para o acréscimo do produto IGC obteríamos o modelo  $DEA_2^*$  da seguinte forma

$$DEA_2^* = \frac{q_1^* TITULADO TOTAL + q_2^* IGC}{DOCENTE TOTAL}$$

As associações apontadas por esse modelo seguem o mesmo padrão do  $DEA_3$ . O modelo  $DEA_3$ , porém, segue a hierarquia dos fatores educacionais identificados na ACP e por isso foi

adotado. Por outro lado, se a escolha do melhor modelo se pautasse em explicar o máximo com o menor número de variáveis possíveis, o modelo  $DEA_2^*$ , provavelmente, seria o contemplado.

### 3.8.1 As instituições eficientes do modelo $DEA_3$

A medida DEA provém de uma aplicação com retornos variáveis à escala voltados à produção, isso porque o objetivo do estudo é maximizar a produção das universidades, que é sintetizada através das variáveis: total de alunos formados na graduação (TITULADO GRD), total de alunos formados na pós-graduação – somando mestrado e doutorado (TITULADO PG) que remetem ao tamanho da universidade e ao IGC que caracteriza a qualidade do ensino fornecido; tal medida varia entre valores maiores ou iguais a um e indica em quanto a universidade precisa multiplicar sua produção sem alteração no consumo de recursos. Como dito na seção anterior, a medida de eficiência é o inverso da medida DEA, de tal maneira, as instituições que demonstrem eficiência continuam com o valor um na medida e as demais com valores inferiores a um e superiores a zero que consiste na proporção entre a produtividade observada e a produtividade máxima (medida DEA igual a um).

Quando uma universidade é dita eficiente não há como melhorar a sua produtividade. Tais universidades serão usadas no modelo para compor a fronteira de eficiências e assim determinar como as demais – ditas ineficientes – devem aumentar sua produção, almejando alcançar tal fronteira.

O valor da medida DEA no modelo orientado para a saída (*output*) indica em quanto à produção de uma universidade deve ser multiplicada, sem se alterar os insumos, para que a mesma seja eficiente. Quando houver folgas, tais valores devem ser somados a produção para obter-se uma meta global. O fato de ser necessária a soma das folgas indica que uma projeção radial da produção não é o suficiente, necessitando haver uma mudança na estrutura da universidade de modo que a desejada eficiência seja alcançada. Embora isso acarrete na variação das proporções da universidade, a alteração no projeto acadêmico é uma maneira de se obter a produção máxima, sendo necessário agregar os valores das folgas como metas mais ambiciosas. Sendo assim, depois de se multiplicar o valor de cada produto pela medida DEA

temos uma meta proporcional; para se alcançar maior status de eficiência é necessária a soma das folgas, resultando assim na meta global.

A Aplicação do modelo resultou em 13 universidades consideradas eficientes. As medidas DEA obtidas com o terceiro modelo estão na primeira coluna da tabela do Apêndice E. Deve ser destacada a situação da Universidade Estadual de Campinas com a medida DEA igual a 1,0041, não sendo considerada eficiente pelo modelo, mas podendo assim ser considerada para qualquer efeito prático.

Nove instituições teriam que ter um aumento proporcional da produção de até 10%, para alcançar a eficiência técnica. O maior grupo encontrado é aquele que contém as universidades que tem que crescer mais de 10% e menos de 20%, contendo 19 participantes. A partir daí os grupos arrefecem, 17 instituições precisam crescer entre 20% e 30%; 11 precisam crescer mais que 30% e menos que 40%; 30 precisam crescer mais de 40% e menos de 70% e por último a Universidade Estadual Feira de Santana terá de dobrar sua produção, crescendo em 103%.

Considerada a projeção radial das instituições ineficientes para a fronteira é necessário investigar as folgas na produção de resultados. As folgas expressam um potencial adicional de crescimento da produção de alguns dos resultados, em proporções distintas, supondo que as instituições possam alterar as proporções com que operam. Elas apontam para mudanças no projeto acadêmico de cada universidade, uma vez que alteram as proporções com que as atividades representadas no conjunto de resultados utilizado têm sido executadas. As folgas obtidas identificam variáveis cujas correspondentes atividades devem ser enfatizadas o que pode acarretar alterações no projeto institucional que dependem de decisões políticas, extrapolando a dimensão da eficiência técnica.

Em relação às folgas encontradas, percebe-se que o maior problema é em relação às titulações da pós-graduação, visto que só há sete universidades com folgas nos titulados da graduação e duas que tem folga no IGC. Além das folgas serem altas para a primeira variável citada a presença delas ocorre em grande parte das universidades, indicando um problema na otimização global da produção. Desse modo, aparentemente, a maior discrepância nas metas globais se daria no número de titulados na pós-graduação, indicando que os projetos institucionais das universidades utilizadas são bastante díspares, o que impacta tal meta tornando-a, em alguns casos, de difícil alcance.

Com o intuito de saber a razão da eficiência das instituições e julgar a viabilidade das metas propostas para as universidades ineficientes é necessária a análise de cada uma das 13 instituições eficientes:

- Universidade Estadual Paulista Júlio de Mesquita Filho (UNIV 25): é caracterizada por possuir o máximo índice de titulação de pós- graduação (valor um), que foi utilizado para se calcular o valor dos índices das demais nessa variável. Percebe-se também que é possuidora de bom IGC, porém baixa formação de estudantes na graduação: 0,2446.
- Universidade do Sagrado Coração (UNIV 29): instituição paulista privada possui baixos índices referentes às conclusões de curso, possuindo o maior para IGC, sendo considerada eficiente apenas sob este aspecto. O fato de seu total titulado na pós-graduação ser quase zero (0,0065), o menor das eficientes, indica que o peso relativo a essa variável é muito menor do que as demais.
- Universidade de Sorocaba (UNIV 30): também privada e paulista, seu maior índice é o referente a IGC (0,5871) e sua situação é similar à Universidade do Sagrado Coração.
- Universidade Estácio de Sá (UNIV 31): universidade carioca e privada, uma das duas eficientes a pertencer ao estado do Rio de Janeiro, demonstra um padrão diferente das demais apresentadas até agora. Seu maior índice é o referente aos concluintes da graduação, com valor de 0,782, o que indica que a eficiência dessa instituição se deve a tal variável. Olhando pelo aspecto dos titulados na pós-graduação, percebe-se que ela é ínfima com índice de 0,0279. Uma nova faceta é utilizada para se projetar suas metas.
- Universidade Cruzeiro do Sul (UNIV 35): universidade paulista e privada, demonstra grandes afinidade com a universidade de Sorocaba, embora todos seus índices sejam maiores os mesmos padrões são encontrados.
- Universidade Nove de Julho (UNIV 38): paulista e privada possui o segundo maior valor de concluintes da graduação, se comparada com as demais eficientes seu conceito de qualidade não é tão alto e seus titulados na pós-graduação são baixos (0,0158).
- Universidade Paulista (UNIV 39): mais uma instituição privada, referência para os concluintes da graduação, possuindo o máximo valor, igual a um. Possui pequena taxa de titulados na pós-graduação, porém relativamente alto IGC. Pode ser caracterizada como o tipo de universidade preocupada em graduar grande quantia de estudantes.

- Universidade Federal do Rio Grande do Sul (UNIV 69): primeira das cinco federais que alcançaram eficiência e a primeira das duas universidades eficientes provenientes do Rio Grande do Sul. Possui o maior valor para IGC, demonstrando ser a universidade com melhor qualidade de curso dos dados utilizados. Percebe-se baixa quantidade de formados na graduação e um nível bom de titulados na pós, se comparado às demais eficientes conteria o segundo maior valor (0,4477). Pode-se inferir que tal universidade tem o objetivo de formar poucos, mas com qualidade, visto que os números da pós são bem menores em termos absolutos.
- Universidade Federal Rural da Amazônia (UNIV 78): universidade federal é a única eficiente que pertence ao Pará. Remete a modelos semelhantes às primeiras instituições eficientes apresentadas, com baixos valores de concluintes e titulações e com considerável IGC (0,6928), formando uma faceta diferenciada.
- Universidade Federal de Lavras (UNIV 80): única instituição mineira tem categoria administrativa pública federal. Seu maior índice é o do IGC, tem semelhanças se comparada à universidade federal do Rio Grande do Sul, tanto que aparecem bastante nas mesmas facetas.
- Fundação Universidade Federal de Ciências da Saúde de Porto Alegre (UNIV 92): universidade gaúcha e pública federal demonstra padrões semelhantes às de Lavras e a UFRGS que possuem apenas altos valores no conceito dos cursos.
- Universidade Estadual do Norte Fluminense Darcy Ribeiro (UNIV 95): instituição pública Estadual possui o menor número de concluintes das eficientes e baixos valores titulados na pós, o alto IGC demonstra que é uma instituição que objetiva formar pouco, mas com qualidade, se outras variáveis fossem levadas em conta poderíamos perceber sua verdadeira razão, podendo ser, por exemplo, voltada a área de pesquisa.
- Fundação Universidade Federal do ABC (UNIV 100): última universidade eficiente possui o segundo melhor IGC (0,9937), segue bastante o padrão da Universidade Estadual do Norte Fluminense Darcy Ribeiro explicada anteriormente.

### 3.8.2 As Facetas obtidas e suas projeções

A fronteira de eficiência estipulada é linear por partes, definida a partir das universidades eficientes de modo que os subconjuntos de referências determinam uma “parede” para as “jaulas” de eficiência nas quais são projetadas todas as outras universidades. A meta global seria a projeção das universidades em tais paredes, melhor chamadas de facetas, nos quais se encontram os melhores valores de produção. De tal maneira, universidades projetadas na mesma faceta demonstrariam uma afinidade de projetos acadêmicos tornando possível assim uma melhor comparação.

As facetas constituem planos de operação mais adequados a cada sistema, proporcionando ações mais tangíveis para redução da ineficiência e aumento da produção.

A medida de eficiência de cada universidade é relativa à distância da universidade à faceta, de modo que universidades longes das projeções receberiam valores menores, o que é mais fácil de visualizar do que na medida DEA, em que valores maiores representariam piores produções.

A viabilidade do alcance das metas depende do aumento proporcional da produção das universidades e de se é possível alterar as especificações dos projetos se a mesma possuir folgas. As características das facetas de eficiência encontradas são analisadas a seguir.

Há dez possíveis facetas principais que demonstram onde as variáveis podem ser projetadas, onde ao menos duas variáveis ineficientes tem suas metas estabelecidas, 14 se considerarmos a projeção de apenas uma instituição ineficiente ou a presença da própria eficiente.

Em ordem decrescente citaremos todas as dez. A primeira é composta por: Fundação Universidade Federal do ABC, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Fundação Universidade Federal de Ciências da Saúde de Porto Alegre e Universidade Estácio de Sá.

A segunda: Fundação Universidade Federal do ABC, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Universidade Federal de Lavras e Universidade Estácio de Sá.

A terceira faceta: Fundação Universidade Federal do ABC, Universidade Federal de Lavras, Universidade Federal do Rio Grande do Sul e Universidade Paulista.



A quarta é formada por: Fundação Universidade Federal do ABC, Universidade Federal de Lavras, Universidade Paulista e Universidade Nove de Julho.

A quinta é formada por: Fundação Universidade Federal do ABC, Universidade Federal de Lavras, Universidade Nove de Julho e Universidade Estácio de Sá.

A sexta e maior: Fundação Universidade Federal do ABC, Universidade Federal do Rio Grande do Sul e Universidade Paulista.

A sétima faceta: Universidade Estadual do Norte Fluminense Darcy Ribeiro, Fundação Universidade Federal de Ciências da Saúde de Porto Alegre, Universidade Cruzeiro do Sul e Universidade Estácio de Sá.

A oitava faceta: Fundação Universidade Federal de Ciências da Saúde de Porto Alegre, Universidade Cruzeiro do Sul, Universidade Estácio de Sá e Universidade de Sorocaba.

A nona faceta é dada por: Universidade Federal de Lavras, Universidade Federal do Rio Grande do Sul, Universidade Paulista e Universidade Nove de Julho.

A décima, e última principal: Universidade Federal de Lavras, Universidade Federal do Rio Grande do Sul, Universidade Nove de Julho e Universidade Estadual Paulista Júlio de Mesquita Filho.

A Tabela 10 resume as universidades eficientes e o número de vezes que aparecem como referência para as demais. Destacam-se as universidades Fundação Universidade Federal do ABC (UNIV100), Universidade Federal do Rio Grande do Sul (UNIV69) e Universidade Paulista (UNIV 39) como as referências mais frequentes. A faceta de eficiência formada por essas instituições é referência para 38 universidades ineficientes dentre estas a maioria das universidades públicas brasileiras.

Tabela 10 - Relação das universidades eficientes com o seu respectivo número de referências

UNIV	Universidade	Número de referências
UNIV 25	Universidade Estadual Júlio de Mesquita Filho	8
UNIV 29	Universidade do Sagrado Coração	1
UNIV 30	Universidade de Sorocaba	6
UNIV 31	Universidade Estácio de Sá	22
UNIV 35	Universidade Cruzeiro do Sul	8
UNIV 38	Universidade Nove de Julho	25
UNIV 39	Universidade Paulista	55
UNIV 69	Universidade Federal do Rio Grande do Sul	57
UNIV 78	Universidade Federal Rural da Amazônia	1
UNIV 80	Universidade Federal de Lavras	33
UNIV 92	Fund. U. Fed. de Ciências da Saúde de Porto Alegre	15
UNIV 95	Universidade Estadual do Norte Fluminense Darcy Ribeiro	11
UNIV 100	Fundação Universidade Federal do ABC	60

A Universidade Federal Rural da Amazônia não é referência para ninguém, podendo entrar numa nova faceta que contenha somente três referências. Podendo ser absorvida tanto na segunda quanto terceira faceta secundárias (que contenha apenas uma instituição ineficiente):

Primeira: Universidade Federal de Lavras, Universidade Nove de Julho, Universidade Estácio de Sá e Universidade Estadual Paulista Júlio de Mesquita Filho.

Segunda: Universidade Federal do Rio Grande do Sul, Universidade Paulista e Universidade Estadual Paulista Júlio de Mesquita Filho.

Terceira: Universidade Cruzeiro do Sul, Universidade Estácio de Sá e Universidade de Sorocaba.

Quarta: Fundação Universidade Federal do ABC, Universidade Federal do Rio Grande do Sul, Universidade Paulista e Universidade Estácio de Sá.

Para efeito de análise, exploraremos como a Universidade de Brasília (UnB-UNIV 2) se comporta referente aos resultados do programa. Tal universidade encontrou um valor DEA muito bom, 1,08, ou seja, para encontrar a sua meta proporcional teria de crescer oito por cento em todas as suas variáveis de produção.

Como os índices (multiplicados por cem), temos valores de 12,49 para o total de concluintes para a graduação, 28,23 para o total de titulados na pós-graduação e 90,61 para o IGC; multiplicando esses valores por 1,08 (o valor DEA) obtemos as metas proporcionais, recebendo os valores constantes da Tabela 11. Somando as folgas, obtemos as metas globais, que resultam na modificação apenas dos titulados, que cresce 2,276, indo para 32,764; sugerindo um aumento maior que o proporcional.

Olhando para a Tabela 12, percebemos os valores absolutos, para as produções da UnB temos os valores 3.735, 1.304 e 3,8811, que seguem para as variáveis Titulado GDR, Titulado PG e IGC. Quando a produção de Titulados PG almeja a meta proporcional temos um aumento de cem unidades, um número bem semelhante à quando se objetiva sair da meta proporcional para a global; um aumento de 300 alunos formados na graduação e do conceito dos cursos em 0,3 também são notados.

Tabela 11 - Resultados para a UnB

	TITULADO GRD	TITULADO PG	IGC
Observado	12,5	28,2	90,6
Meta proporcional	13,5	30,5	97,8
Folgas	0,00	2,3	0,00
Meta global	13,5	32,8	97,8

Tabela 12 - Resultados para a UnB em valores originais

	TITULADO GRD	TITULADO PG	IGC
Observado	3.735	1.304	3,8
Meta proporcional	4.033,8	1.408,3	4,2
Folgas	0,00	105,1	0,00
Meta global	4.033,8	1.513,4	4,2

Se observarmos a faceta na qual a UnB foi projetada, descobrimos que ela pertence à maior delas, composta pelas universidades: Fundação Universidade Federal do ABC (UNIV 100), Universidade Federal do Rio Grande do Sul (UNIV 69) e Universidade Paulista (UNIV 39); nas proporções: 0,2125; 0,725 e 0,0623. Ou seja, para se definir as metas globais para as outras variáveis propostas não inclusas no modelo DEA, basta se multiplicar os valores pelas variáveis das outras universidades, no seguinte esquema:

$$Meta(UNIV 2) = (UNIV 100).0,212 + (UNIV 69).0,725 + (UNIV 39).0,062$$

Mas por que nessa faceta? Observando-se os valores absolutos das variáveis de produção das universidades percebe-se que tanto a Universidade Federal do Rio Grande do Sul quanto a Universidade Federal do ABC apresentam grandes valores para os titulados da pós-graduação o que explica o fato de ser a única variável com folga na análise do desempenho da UnB. Em relação ao IGC tanto a do ABC quanto a Universidade do Rio Grande do Sul se mostram como referências. A Universidade Paulista entra como corretora de porte, tanto no número de concluintes quanto no de total de docentes.

Como se pode perceber no Apêndice F, a divisão de cada produto pelo único insumo (Total de docentes) proporciona uma maneira mais viável de se perceber o porquê das referências. Em relação ao indicador Titulados da graduação por docente, percebe-se que os maiores valores pertencem às universidades privadas, aquelas que têm o objetivo de formar bastante com um professor com o intuito de manter alto fluxo de estudantes, universidades como as UNIV'S 31, 35 e 38 estão no topo do ranking.

Na variável de titulados da pós por docente temos universidades federais e estaduais públicas no topo do ranking, as UNIV'S eficientes 25, 95, 69 e 80 aparecem no topo; instituições que tem objetivos de instruir de maneira altamente capacitante seus alunos.

A variável de IGC por docente apresenta valores bem reduzidos, isso se deve ao fato do valor do IGC ser bem inferior ao do total de docentes, no topo se encontram universidades públicas e privadas, não mostrando um padrão certo, as UNIV'S 92, 95, 78, 29 e 100 encabeçam a lista de maiores IGC por docente, ou seja, conseguem uma nota de curso maior com menos professores.

A Universidade Paulista possui grande número de titulados na graduação, de modo que sugere uma boa indicação para essa variável, 3,35 alunos graduados por docente. A Universidade Federal do Rio Grande do Sul é uma boa referência para a segunda variável, número de titulados na pós-graduação, mostrando que forma 0,8047 alunos (em mestrado ou doutorado) para cada docente. Em relação a última variável, a de conceito das instituições, a Universidade Federal do ABC se mostra bem eficiente em relação à mesma, com incríveis 0,0096 pontos de IGC por docente, o maior encontrado nos dados, explicando a presença na faceta.

Em resumo, foram identificadas 13 universidades eficientes que compõe a fronteira. A principal faceta de eficiência identificada é formada pelas instituições: Universidade Federal

do Rio Grande do Sul, Fundação Universidade Federal do ABC e Universidade Paulista que constituem um espaço de eficiência caracterizado pela alta produção da pós-graduação e valores elevados no IGC. Os resultados obtidos para a Universidade de Brasília foram utilizados para ilustrar o procedimento de produção para a fronteira e a identificação de ações e estratégias que conduzem ao cumprimento das metas propostas. O Apêndice G traz a relação das folgas e metas proporcionais e globais das demais universidades.

## CONCLUSÕES E RECOMENDAÇÕES

Esta pesquisa utilizou uma combinação de Análise Estatística Exploratória de Dados e Análise Envoltória de Dados para desenvolver uma metodologia de avaliação da eficiência produtiva das universidades brasileiras.

A análise estatística exploratória dos dados baseou-se em medidas estatísticas simples, análise de correlações lineares e análise em componentes principais e permitiu a descrição das universidades e a análise das variáveis e suas inter-relações possibilitando a identificação dos principais fatores educacionais presentes no banco de dados, seus descritores e a viabilidade de uso em um modelo DEA.

O modelo DEA com retornos variáveis e orientado para a maximização da produção mostrou-se eficaz na formulação de metas que conduzam à melhoria da produção universitária. Tal modelo foi capaz de definir os diferentes tipos de eficiências das instituições levando em conta suas individualidades e projetos acadêmicos, sugerindo ações e estratégias para o alcance das metas propostas.

Esta pesquisa não considerou a imposição de restrições sobre os multiplicadores (pesos) do modelo DEA além das restrições típicas da técnica. O uso adequado de restrições aos pesos poderia identificar melhor os projetos das instituições de pequeno porte bem como evitar a construção de facetos específicas para uma ou duas instituições. A investigação de modelos com restrição nos pesos poderá conduzir a uma melhor definição dos projetos eficientes e, em consequência, um melhor detalhamento da fronteira de eficiência.

## REFERÊNCIAS

- BANKER, R.D.; CHARNES A.; COOPER, W.W. Some models for estimation technical and scale inefficiencies in Data Envelopment Analysis. **Management Science**, 30(9):1078-1092, 1984.
- BELLONI, J.A. **Curso de especialização em estatística aplica: Análise em componentes principais**. Universidade de Brasília, março de 2004.
- BELLONI, J.A. **Uma Metodologia de Avaliação da Eficiência Produtiva de Universidades Federais Brasileiras**. Florianópolis, 2000. Tese de Doutorado, UFSC, 2000.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Senado, 1988.
- BRASIL. Portaria Normativa nº40, de 12 de dezembro de 2007. **Instituição do e-MEC**. Disponível em: <<http://meclegis.mec.gov.br/documento/view/id/17>>. Acesso em: 15 Set.2013.
- BRASIL. Decreto nº5773, de maio de 2006. **Dispõe sobre o exercício das funções de regulação, supervisão e avaliação de instituições de educação superior e cursos superiores de graduação e sequenciais no sistema federal de ensino**. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2006/decreto/d5773.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/decreto/d5773.htm)>. Acesso em: 27 Set.2013.
- BRASIL. Lei nº11.892, de 29 de dezembro de 2008. **Institui a Rede Federal de Educação Profissional, Científica e Tecnológica, cria os Institutos Federais de Educação, Ciência e Tecnologia, e dá outras providências**. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2007-2010/2008/lei/111892.htm](http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/111892.htm)>. Acesso em: 27 Set.2013.
- BRASIL. Lei nº11.892, de 29 de dezembro de 2008. **Transforma o Instituto Nacional de Estudos e Pesquisas Educacionais – INEP em autarquia federal e dá outras providências**. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/Leis/L9448.htm](http://www.planalto.gov.br/ccivil_03/Leis/L9448.htm)>. Acesso em: 27 Set.2013.

BRASIL. Decreto nº 6425, de 4 de abril de 2008. **Dispõe sobre o censo anual da educação.** Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2007-2010/2008/Decreto/D6425.htm](http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/Decreto/D6425.htm)>. Acesso em: 27 Set.2013.

BRASIL. **Modifica as competências e a estrutura organizacional da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.** Disponível em: < [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2007-2010/2007/Lei/L11502.htm](http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Lei/L11502.htm) >. Acesso em: 30 Set.2013.

CARVALHO FERREIRA, C.M. e GOMES, A. P. **Introdução à Análise Envoltória de Dados – Teoria, modelos e aplicações**, FGV, 2009.

CAPES. **Distribuição de discentes de pós-graduação no Brasil.** Disponível em: < <http://geocapes.capes.gov.br/geocapesds/#app=c501&da7a-selectedIndex=0&5317-selectedIndex=0&dbcb-selectedIndex=0>>. Acesso em: 15 Out.2013.

CHARNES, A.; COOPER, W. W.; LEWIN, A. Y.; SEIFORD, L. M. **Data Envelopment Analysis: Theory, Methodology and Application.** 2 Ed., Kluwer Academic Publishers, Boston, 1993.

CHARNES, A.; COOPER, W. W. Preface to topics in Data Envelopment Analysis. **Annals of Operations Research**, 2, 59-94, 1985.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, 2(6),429-444, 1978.

ESTELLITA LINS, M. A. P. E ÂNGULO MEZA, L. (cords.). **Análise Envoltória de Dados**, COPPE-UFRJ, 2000.

FARREL, M.J. The measurement of productive efficiency. **J. Royal Statistical Society**, v. 120, Part III, 253-290, 1957.

Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. **Journal of Educational Psychology**, 24, 417-441.

INEP. **Censo da Educação Superior.** Disponível em: <<http://portal.inep.gov.br/web/censo-da-educacao-superior>>. Acesso em: 03 Ago.2013.

INEP. **Resumo técnico censo da Educação Superior 2010.** Disponível em: < <http://portal.inep.gov.br/resumos-tecnicos>>. Acesso em: 04 Ago.2013.



- INEP. **Resumo técnico censo da Educação Superior 2011**. Disponível em: <<http://portal.inep.gov.br/resumos-tecnicos>>. Acesso em: 04 Ago.2013.
- JOHNSON, Richard Arnold; WICHERN, Dean W. **Applied multivariate statistical analysis**. 6th ed. Upper Saddle River: Pearson Prentice Hall, c2007. xviii, 773 p.
- NORMAN, M. e STOKER, B. **Data Envelopment Analysis: The Assesment of Performance**. John Wiley & Sons, 1991.
- PORTAL BRASIL. **Ensino superior**. Disponível em <<http://www.brasil.gov.br/educacao/2009/11/ensino-superior>>. Acesso em 09 Out.2013.
- RENCHER, Alvin C. **Methods of multivariate analysis**. 2nd ed. New York: J. Wiley, c2002. xxii, 708 p.
- RHODES, E.L. **Data Envelopment Analysis and approaches for measuring the efficiency of decision making units with an application to program follow-through in U.S. education**. Ph.D. Dissertation, Carnegie Mellon University, 1978.
- SHARMA, Subhash. **Applied multivariate techniques**. New york: John Wiley & Sons, 1996. 493 p.

## **APÊNDICE A – Instituições da educação superior**

A Classificação das Instituições de Ensino Superior (IES), quanto a sua organização acadêmica, utilizada neste trabalho, é aquela amparada principalmente pela Portaria Normativa nº40, de 12 de dezembro de 2007, Com o auxílio de outras normas infraconstitucionais quando necessário, sendo definidas como a seguir.

### **Faculdades**

São institutos e organizações equiparadas, nos termos do Decreto nº5773, de maio de 2006. Caracterizadas por não cobrirem todas as áreas do conhecimento humano e não realizarem pesquisa.

### **Centros Universitários**

Instituições de ensino superior pluricurriculares que se caracterizam pela excelência no ensino oferecido, dotado de autonomia para a criação de cursos e vagas na sede, está obrigado a manter um terço de mestres ou doutores e um quinto do corpo docente em tempo integral.

### **Universidades**

Instituições pluricurriculares de formação dos quadros profissionais de nível superior, de pesquisa, de extensão e de domínio e cultivo do saber humano. São dotadas de autonomia na sede, podem criar campus fora de sede no âmbito do Estado e estão obrigadas a manter um terço de mestres ou doutores e um terço do corpo docente em tempo integral. Tem como fim o ensino, a pesquisa e a extensão em todas as áreas do conhecimento humano.

### **Institutos federais de educação, ciência e tecnologia**

São instituições de educação superior, básica e profissional, pluricurriculares e multicampi, especializados na oferta de educação profissional e tecnológica nas diferentes modalidades de ensino. Para efeitos regulatórios equiparam-se às universidades federais (Lei nº11.892, de 29 de dezembro de 2008).

### **Centros federais de educação tecnológica**

Tem por finalidade formar e qualificar profissionais nos diferentes níveis e modalidades de ensino, para os diversos setores da economia, além de realizar pesquisas aplicadas e promover o desenvolvimento tecnológico. Para efeitos regulatórios equiparam-se aos centros universitários.

### APÊNDICE B – A matriz de correlações

Variáveis	MATRICULA GRD	MATRICULA GRD PRESC	VAGA TOTAL	VAGA PRESC	ING TOTAL	ING PRESC	TITULADO GRD	TITULADO PRESC
MATRICULA GRD	1,00	0,98	0,89	0,93	0,97	0,96	0,96	0,97
MATRICULA GRD PRESC	0,98	1,00	0,83	0,90	0,93	0,96	0,93	0,98
VAGA TOTAL	0,89	0,83	1,00	0,98	0,91	0,90	0,83	0,85
VAGA PRESC	0,93	0,90	0,98	1,00	0,95	0,96	0,88	0,91
ING TOTAL	0,97	0,93	0,91	0,95	1,00	0,97	0,96	0,93
ING PRESC	0,96	0,96	0,90	0,96	0,97	1,00	0,91	0,93
TITULADO GRD	0,96	0,93	0,83	0,88	0,96	0,91	1,00	0,95
TITULADO PRESC	0,97	0,98	0,85	0,91	0,93	0,93	0,95	1,00
DOC OUTROS	0,89	0,86	0,93	0,94	0,89	0,89	0,82	0,85
DOC MESTRADO	0,92	0,90	0,90	0,91	0,89	0,90	0,84	0,87
DOC DOUTORADO	0,42	0,45	0,36	0,35	0,35	0,38	0,34	0,39
DOCENTE TOTAL	0,83	0,83	0,78	0,79	0,77	0,79	0,74	0,78
QT ACERVO	0,58	0,58	0,50	0,50	0,51	0,50	0,54	0,57
CONCEITO GRD	0,09	0,08	0,18	0,13	0,10	0,12	0,00	0,02
CONCEITO MESTRADO	0,17	0,19	0,13	0,13	0,12	0,14	0,15	0,19
CONCEITO DOUTORADO	0,14	0,17	0,10	0,10	0,10	0,13	0,10	0,14
IGC	0,02	0,02	0,09	0,04	0,01	0,04	-0,07	-0,05
MATRICULA MESTRADO	0,38	0,39	0,37	0,35	0,33	0,34	0,33	0,37
TITULADO MESTRADO	0,38	0,39	0,36	0,34	0,32	0,34	0,33	0,37
MATRICULA DOUTORADO	0,32	0,34	0,29	0,28	0,27	0,29	0,28	0,31
TITULADO DOUTORADO	0,30	0,32	0,27	0,26	0,25	0,27	0,26	0,29
MATRICULA PG	0,36	0,37	0,34	0,33	0,31	0,33	0,31	0,35
TITULADO PG	0,36	0,38	0,34	0,33	0,31	0,33	0,32	0,35
MATRICULA TOTAL	0,97	1,00	0,82	0,89	0,91	0,95	0,92	0,97
TITULADO TOTAL	0,97	0,97	0,82	0,89	0,91	0,92	0,94	0,99

Variáveis	DOC OUTROS	DOC MESTRADO	DOC DOUTORADO	DOCENTE TOTAL	QT ACERVO
MATRICULA GRD	0,89	0,92	0,42	0,83	0,58
MATRICULA GRD PRESC	0,86	0,90	0,45	0,83	0,58
VAGA TOTAL	0,93	0,90	0,36	0,78	0,50
VAGA PRESC	0,94	0,91	0,35	0,79	0,50
ING TOTAL	0,89	0,89	0,35	0,77	0,51
ING PRESC	0,89	0,90	0,38	0,79	0,50
TITULADO GRD	0,82	0,84	0,34	0,74	0,54
TITULADO PRESC	0,85	0,87	0,39	0,78	0,57
DOC OUTROS	1,00	0,94	0,40	0,86	0,48
DOC MESTRADO	0,94	1,00	0,44	0,88	0,53
DOC DOUTORADO	0,40	0,44	1,00	0,79	0,70
DOCENTE TOTAL	0,86	0,88	0,79	1,00	0,69
QT ACERVO	0,48	0,53	0,70	0,69	1,00
CONCEITO GRD	0,07	0,12	0,40	0,25	0,22
CONCEITO MESTRADO	0,10	0,09	0,53	0,32	0,45
CONCEITO DOUTORADO	0,09	0,10	0,53	0,32	0,44
IGC	0,00	0,03	0,59	0,30	0,30
MATRICULA MESTRADO	0,36	0,37	0,88	0,66	0,68
TITULADO MESTRADO	0,35	0,36	0,89	0,66	0,68
MATRICULA DOUTORADO	0,28	0,29	0,91	0,63	0,72
TITULADO DOUTORADO	0,26	0,27	0,90	0,61	0,71
MATRICULA PG	0,33	0,34	0,90	0,66	0,70
TITULADO PG	0,33	0,35	0,90	0,66	0,69
MATRICULA TOTAL	0,85	0,89	0,51	0,85	0,61
TITULADO TOTAL	0,83	0,86	0,48	0,82	0,62

Variáveis	CONCEITO GRD	CONCEITO MESTRADO	CONCEITO DOUTORADO	IGC
MATRICULA GRD	0,09	0,17	0,14	0,02
MATRICULA GRD PRESC	0,08	0,19	0,17	0,02
VAGA TOTAL	0,18	0,13	0,10	0,09
VAGA PRESC	0,13	0,13	0,10	0,04
ING TOTAL	0,10	0,12	0,10	0,01
ING PRESC	0,12	0,14	0,13	0,04
TITULADO GRD	0,00	0,15	0,10	-0,07
TITULADO PRESC	0,02	0,19	0,14	-0,05
DOC OUTROS	0,07	0,10	0,09	0,00
DOC MESTRADO	0,12	0,09	0,10	0,03
DOC DOUTORADO	0,40	0,53	0,53	0,59
DOCENTE TOTAL	0,25	0,32	0,32	0,30
QT ACERVO	0,22	0,45	0,44	0,30
CONCEITO GRD	1,00	0,26	0,22	0,89
CONCEITO MESTRADO	0,26	1,00	0,83	0,47
CONCEITO DOUTORADO	0,22	0,83	1,00	0,45
IGC	0,89	0,47	0,45	1,00
MATRICULA MESTRADO	0,36	0,49	0,48	0,50
TITULADO MESTRADO	0,36	0,52	0,50	0,52
MATRICULA DOUTORADO	0,38	0,60	0,58	0,57
TITULADO DOUTORADO	0,38	0,62	0,61	0,59
MATRICULA PG	0,37	0,54	0,52	0,53
TITULADO PG	0,37	0,55	0,53	0,54
MATRICULA TOTAL	0,10	0,23	0,20	0,06
TITULADO TOTAL	0,05	0,25	0,20	0,02

Variáveis	MATRICULA MESTRADO	TITULADO MESTRADO	MATRICULA DOUTORADO	TITULADO DOUTORADO	MATRICULA PG	TITULADO PG
MATRICULA GRD	0,38	0,38	0,32	0,30	0,36	0,36
MATRICULA GRD PRESC	0,39	0,39	0,34	0,32	0,37	0,38
VAGA TOTAL	0,37	0,36	0,29	0,27	0,34	0,34
VAGA PRESC	0,35	0,34	0,28	0,26	0,33	0,33
ING TOTAL	0,33	0,32	0,27	0,25	0,31	0,31
ING PRESC	0,34	0,34	0,29	0,27	0,33	0,33
TITULADO GRD	0,33	0,33	0,28	0,26	0,31	0,32
TITULADO PRESC	0,37	0,37	0,31	0,29	0,35	0,35
DOC OUTROS	0,36	0,35	0,28	0,26	0,33	0,33
DOC MESTRADO	0,37	0,36	0,29	0,27	0,34	0,35
DOC DOUTORADO	0,88	0,89	0,91	0,90	0,90	0,90
DOCENTE TOTAL	0,66	0,66	0,63	0,61	0,66	0,66
QT ACERVO	0,68	0,68	0,72	0,71	0,70	0,69
CONCEITO GRD	0,36	0,36	0,38	0,38	0,37	0,37
CONCEITO MESTRADO	0,49	0,52	0,60	0,62	0,54	0,55
CONCEITO DOUTORADO	0,48	0,50	0,58	0,61	0,52	0,53
IGC	0,50	0,52	0,57	0,59	0,53	0,54
MATRICULA MESTRADO	1,00	1,00	0,96	0,93	0,99	0,99
TITULADO MESTRADO	1,00	1,00	0,96	0,94	0,99	1,00
MATRICULA DOUTORADO	0,96	0,96	1,00	0,99	0,98	0,98
TITULADO DOUTORADO	0,93	0,94	0,99	1,00	0,96	0,96
MATRICULA PG	0,99	0,99	0,98	0,96	1,00	1,00
TITULADO PG	0,99	1,00	0,98	0,96	1,00	1,00
MATRICULA TOTAL	0,45	0,45	0,41	0,38	0,44	0,44
TITULADO TOTAL	0,46	0,46	0,41	0,39	0,44	0,45

Variáveis	MATRICULA TOTAL	TITULADO TOTAL
MATRICULA GRD	0,97	0,97
MATRICULA GRD PRESC	1,00	0,97
VAGA TOTAL	0,82	0,82
VAGA PRESC	0,89	0,89
ING TOTAL	0,91	0,91
ING PRESC	0,95	0,92
TITULADO GRD	0,92	0,94
TITULADO PRESC	0,97	0,99
DOC OUTROS	0,85	0,83
DOC MESTRADO	0,89	0,86
DOC DOUTORADO	0,51	0,48
DOCENTE TOTAL	0,85	0,82
QT ACERVO	0,61	0,62
CONCEITO GRD	0,10	0,05
CONCEITO MESTRADO	0,23	0,25
CONCEITO DOUTORADO	0,20	0,20
IGC	0,06	0,02
MATRICULA MESTRADO	0,45	0,46
TITULADO MESTRADO	0,45	0,46
MATRICULA DOUTORADO	0,41	0,41
TITULADO DOUTORADO	0,38	0,39
MATRICULA PG	0,44	0,44
TITULADO PG	0,44	0,45
MATRICULA TOTAL	1,00	0,98
TITULADO TOTAL	0,98	1,00



### APÊNDICE C - Relação das instituições utilizadas

UNIV	UNIVERSIDADE	CATEGORIA ADMINISTRATIVA	UF
1	UNIVERSIDADE FEDERAL DE MATO GROSSO	Publica Federal	MT
2	UNIVERSIDADE DE BRASÍLIA	Publica Federal	DF
3	UNIVERSIDADE FEDERAL DE SERGIPE	Publica Federal	SE
4	UNIVERSIDADE FEDERAL DO AMAZONAS	Publica Federal	AM
5	UNIVERSIDADE FEDERAL DO PIAUÍ	Publica Federal	PI
6	UNIVERSIDADE FEDERAL DE OURO PRETO	Publica Federal	MG
7	UNIVERSIDADE FEDERAL DE SÃO CARLOS	Publica Federal	SP
8	UNIVERSIDADE FEDERAL DE VIÇOSA	Publica Federal	MG
9	UNIVERSIDADE ESTADUAL DE LONDRINA	Publica Estadual	PR
10	PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ	Privada	PR
11	UNIVERSIDADE FEDERAL DO RIO GRANDE	Publica Federal	RS
12	UNIVERSIDADE DE CAXIAS DO SUL	Privada	RS
13	UNIVERSIDADE DO VALE DO RIO DOS SINOS	Privada	RS
14	UNIVERSIDADE GAMA FILHO	Privada	RJ
15	UNIVERSIDADE FEDERAL DE UBERLÂNDIA	Publica Federal	MG
16	UNIVERSIDADE CATÓLICA DE PELOTAS	Privada	RS
17	PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS	Privada	SP
18	UNIVERSIDADE DE PASSO FUNDO	Privada	RS
19	PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL	Privada	RS
20	UNIVERSIDADE PRESBITERIANA MACKENZIE	Privada	SP
21	UNIVERSIDADE ESTADUAL DE SANTA CRUZ	Publica Estadual	BA
22	UNIVERSIDADE ESTADUAL DO CEARÁ	Publica Estadual	CE
23	FUNDAÇÃO UNIVERSIDADE DO ESTADO DE SANTA CATARINA	Publica Estadual	SC
24	UNIVERSIDADE ESTADUAL DE CAMPINAS	Publica Estadual	SP
25	UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO	Publica Estadual	SP
26	UNIVERSIDADE ESTADUAL DE MARINGÁ	Publica Estadual	PR
27	UNIVERSIDADE REGIONAL DE BLUMENAU	Publica Municipal	SC
28	UNIVERSIDADE DO VALE DO ITAJAÍ	Privada	SC
29	UNIVERSIDADE DO SAGRADO CORAÇÃO	Privada	SP
30	UNIVERSIDADE DE SOROCABA	Privada	SP
31	UNIVERSIDADE ESTÁCIO DE SÁ	Privada	RJ
32	UNIVERSIDADE METODISTA DE SÃO PAULO	Privada	SP
33	UNIVERSIDADE SÃO JUDAS TADEU	Privada	SP
34	UNIVERSIDADE DE RIBEIRÃO PRETO	Privada	SP
35	UNIVERSIDADE CRUZEIRO DO SUL	Privada	SP
36	UNIVERSIDADE METODISTA DE PIRACICABA	Privada	SP
37	UNIVERSIDADE DE SANTA CRUZ DO SUL	Privada	RS
38	UNIVERSIDADE NOVE DE JULHO	Privada	SP

UNIV	UNIVERSIDADE	CATEGORIA ADMINISTRATIVA	UF
39	UNIVERSIDADE PAULISTA	Privada	SP
40	PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS	Privada	MG
41	UNIVERSIDADE TUIUTI DO PARANÁ	Privada	PR
42	UNIVERSIDADE SALVADOR	Privada	BA
43	UNIVERSIDADE CATÓLICA DE BRASÍLIA	Privada	DF
44	UNIVERSIDADE REGIONAL INTEGRADA DO ALTO URUGUAI E DAS MISSÕES	Privada	RS
45	UNIVERSIDADE LUTERANA DO BRASIL	Privada	RS
46	UNIVERSIDADE BANDEIRANTE DE SÃO PAULO	Privada	SP
47	UNIVERSIDADE GUARULHOS	Privada	SP
48	UNIVERSIDADE DO EXTREMO SUL CATARINENSE	Privada	SC
49	UNIVERSIDADE DE FRANCA	Privada	SP
50	UNIVERSIDADE CATÓLICA DO SALVADOR	Privada	BA
51	UNIVERSIDADE DE MOGI DAS CRUZES	Privada	SP
52	PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS	Privada	GO
53	PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO	Privada	RJ
54	PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO	Privada	SP
55	UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO	Publica Estadual	RJ
56	UNIVERSIDADE FEDERAL DO MARANHÃO	Publica Federal	MA
57	UNIVERSIDADE FEDERAL DO PARÁ	Publica Federal	PA
58	UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE	Publica Federal	RN
59	UNIVERSIDADE FEDERAL DO PARANÁ	Publica Federal	PR
60	UNIVERSIDADE FEDERAL FLUMINENSE	Publica Federal	RJ
61	UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO	Publica Federal	ES
62	UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO	Publica Federal	RJ
63	UNIVERSIDADE FEDERAL DE MINAS GERAIS	Publica Federal	MG
64	UNIVERSIDADE FEDERAL DE JUIZ DE FORA	Publica Federal	MG
65	UNIVERSIDADE FEDERAL DE ALAGOAS	Publica Federal	AL
66	UNIVERSIDADE FEDERAL DA BAHIA	Publica Federal	BA
67	UNIVERSIDADE FEDERAL DA PARAÍBA	Publica Federal	PB
68	UNIVERSIDADE FEDERAL DE PERNAMBUCO	Publica Federal	PE
69	UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL	Publica Federal	RS
70	UNIVERSIDADE FEDERAL DE SANTA MARIA	Publica Federal	RS
71	UNIVERSIDADE FEDERAL DO CEARÁ	Publica Federal	CE
72	UNIVERSIDADE FEDERAL DE GOIÁS	Publica Federal	GO
73	UNIVERSIDADE FEDERAL DE SANTA CATARINA	Publica Federal	SC
74	UNIVERSIDADE FEDERAL DO RIO DE JANEIRO	Publica Federal	RJ
75	UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO	Publica Federal	PE
76	UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ	Publica Federal	PR
77	UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO	Publica Federal	RN
78	UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA	Publica Federal	PA
79	UNIVERSIDADE FEDERAL DE SÃO PAULO	Publica Federal	SP
80	UNIVERSIDADE FEDERAL DE LAVRAS	Publica Federal	MG
81	UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO	Publica Federal	MG
82	UNIVERSIDADE FEDERAL DE ITAJUBÁ – UNIFEI	Publica Federal	MG

UNIV	UNIVERSIDADE	CATEGORIA ADMINISTRATIVA	UF
83	UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ	Publica Estadual	PR
84	UNIVERSIDADE FEDERAL DE PELOTAS	Publica Federal	RS
85	UNIVERSIDADE DE TAUBATÉ	Publica Municipal	SP
86	UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA	Publica Estadual	BA
87	UNIVERSIDA SÃO FRANCISCO	Privada	SP
88	UNIVERSIDADE ESTADUAL DO SUDOESTE DA BAHIA	Publica Estadual	BA
89	UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO	Publica Federal	RJ
90	UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL	Publica Federal	MS
91	FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA	Publica Federal	RO
92	FUNDAÇÃO UNIVERSIDADE FEDERAL DE CIÊNCIAS DA SAÚDE DE PORTO ALEGRE	Publica Federal	RS
93	UNIVERSIDADE FEDERAL DO AMAPÁ	Publica Federal	AP
94	UNIVERSIDADE METROPOLITANA DE SANTOS	Privada	SP
95	UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY RIBEIRO	Publica Estadual	RJ
96	UNIVERSIDADE FEDERAL DE CAMPINA GRANDE	Publica Federal	PB
97	UNIVERSIDADE DO ESTADO DO AMAZONAS	Publica Estadual	AM
98	UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA	Publica Federal	BA
99	FUNDAÇÃO UNIVERSIDADE FEDERAL DA GRANDE DOURADOS	Publica Federal	MS
100	FUNDAÇÃO UNIVERSIDADE FEDERAL DO ABC	Publica Federal	SP

**APÊNDICE D – Correlações das variáveis com os índices de eficiência**

Variáveis	EFIC1	EFIC2	EFIC3	EFIC2*
MATRICULA GRD	0,370	0,319	0,234	0,238
MATRICULA GRD PRESC	0,358	0,319	0,259	0,264
VAGA TOTAL	0,352	0,290	0,185	0,189
VAGA PRESC	0,429	0,351	0,226	0,230
ING TOTAL	0,401	0,326	0,216	0,220
ING PRESC	0,394	0,330	0,260	0,265
TITULADO GRD	0,462	0,393	0,213	0,215
TITULADO PRESC	0,480	0,426	0,277	0,281
DOC OUTROS	0,219	0,106	0,028	0,032
DOC MESTRADO	0,145	0,038	0,029	0,034
DOC DOUTORADO	-0,298	0,020	0,389	0,386
DOCENTE TOTAL	-0,045	0,062	0,246	0,247
QT ACERVO	0,042	0,223	0,302	0,301
CONCEITO GRD	-0,369	-0,097	0,603	0,609
CONCEITO MESTRADO	-0,082	0,227	0,383	0,383
CONCEITO DOUTORADO	-0,261	0,024	0,240	0,240
IGC	-0,467	-0,066	0,648	0,654
MATRICULA MESTRADO	-0,230	0,128	0,342	0,333
TITULADO MESTRADO	-0,227	0,146	0,360	0,353
MATRICULA DOUTORADO	-0,215	0,169	0,415	0,410
TITULADO DOUTORADO	-0,220	0,178	0,425	0,422
MATRICULA PG	-0,227	0,147	0,377	0,370
TITULADO PG	-0,228	0,156	0,381	0,374
MATRICULA TOTAL	0,330	0,328	0,291	0,296
TITULADO TOTAL	0,435	0,439	0,328	0,330
EFIC1	1,000	0,245	-0,480	-0,473
EFIC2	0,788	1,000	0,079	0,081
EFIC3	-0,480	0,079	1,000	1,000
EFIC2*	-0,473	0,081	1,000	1,000

**APÊNDICE E – Valores observados e folgas**

UNIV	DEA	Observado			Folgas		
		TITULADO GRD	TITULADO PG	IGC	TITULADO GRD	TITULADO PG	IGC
1	1,37	6,900	8,360	71,800	0,000	12,038	0,000
2	1,08	12,490	28,230	90,610	0,000	2,276	0,000
3	1,40	6,620	7,750	69,910	0,000	0,000	0,000
4	1,56	6,810	7,340	62,720	0,000	6,265	0,000
5	1,40	8,550	5,630	69,610	0,000	3,181	0,000
6	1,25	2,360	5,410	79,670	0,100	2,498	0,000
7	1,06	3,780	16,300	93,750	0,524	0,065	0,000
8	1,03	5,420	18,120	95,180	0,000	0,000	0,000
9	1,21	7,770	10,370	81,770	0,000	13,321	0,000
10	1,24	22,810	6,540	67,530	0,000	0,000	0,000
11	1,34	3,020	4,740	74,010	0,000	0,173	0,000
12	1,35	11,660	2,340	67,890	0,000	0,000	0,000
13	1,19	7,210	7,730	80,750	0,000	0,000	0,000
14	1,67	6,630	0,950	57,010	0,000	0,000	0,000
15	1,18	8,500	12,320	83,300	0,000	3,082	0,000
16	1,22	2,840	1,470	69,800	0,000	0,000	0,000
17	1,49	10,290	1,100	61,340	0,000	0,000	0,000
18	1,39	8,250	2,490	68,350	0,000	0,000	0,000
19	1,10	12,380	14,510	85,910	0,000	0,000	0,000
20	1,29	18,270	5,720	67,770	0,000	0,000	0,000
21	1,41	2,510	2,380	70,300	0,000	5,036	0,000
22	1,52	6,060	7,840	63,700	0,000	0,000	0,000
23	1,14	4,740	4,870	86,130	0,000	0,000	0,000
24	1,00	8,350	32,970	98,410	0,000	0,000	0,000
25	1,00	24,460	100,000	87,660	0,000	0,000	0,000
26	1,18	7,000	12,710	83,960	0,000	7,472	0,000
27	1,48	5,120	2,920	66,300	0,000	0,000	0,000
28	1,42	10,320	4,010	66,720	0,000	0,000	0,000
29	1,00	2,460	0,650	71,190	0,000	0,000	0,000
30	1,00	5,420	1,000	58,710	0,000	0,000	0,000
31	1,00	78,200	2,790	48,930	0,000	0,000	0,000
32	1,21	9,460	3,380	59,320	0,000	0,000	6,978
33	1,12	10,080	0,690	63,050	0,000	0,445	0,000
34	1,34	4,100	0,500	68,210	0,000	0,626	0,000
35	1,00	10,140	2,880	68,050	0,000	0,000	0,000
36	1,54	5,850	2,880	58,010	0,000	0,000	0,000
37	1,36	4,070	2,230	71,830	0,000	0,000	0,000
38	1,00	64,670	1,580	64,630	0,000	0,000	0,000
39	1,00	100,000	1,710	73,570	0,000	0,000	0,000
40	1,25	25,240	8,720	66,760	0,000	0,000	0,000

UNIV	DEA	Observado			Folgas		
		TITULADO GRD	TITULADO PG	IGC	TITULADO GRD	TITULADO PG	IGC
41	1,58	5,090	1,390	55,620	0,000	0,000	0,000
42	1,41	4,850	0,870	62,700	0,000	0,010	0,000
43	1,43	7,310	4,330	66,320	0,000	0,000	0,000
44	1,54	5,220	1,230	63,080	0,000	0,000	0,000
45	1,46	18,340	1,950	57,430	0,000	0,000	0,000
46	1,13	28,130	1,710	43,640	0,000	0,000	7,419
47	1,14	13,490	0,630	50,960	0,000	0,534	0,000
48	1,56	5,380	1,260	61,380	0,000	0,000	0,000
49	1,49	6,100	1,210	55,330	0,000	0,000	0,000
50	1,19	6,750	1,190	58,740	0,000	0,000	0,000
51	1,53	9,920	0,670	55,190	0,000	0,416	0,000
52	1,60	20,540	4,310	51,330	0,000	0,000	0,000
53	1,12	6,520	14,310	88,110	0,000	0,000	0,000
54	1,11	8,830	25,960	87,510	0,000	0,000	0,000
55	1,41	7,720	20,390	70,770	0,000	15,465	0,000
56	1,43	5,270	4,550	69,100	0,000	9,676	0,000
57	1,40	12,400	18,080	69,210	0,000	7,089	0,000
58	1,16	9,810	17,770	85,450	0,000	21,338	0,000
59	1,20	8,270	26,800	82,730	0,000	8,877	0,000
60	1,25	12,040	20,390	78,860	0,000	16,740	0,000
61	1,34	6,420	12,540	73,930	0,000	3,445	0,000
62	1,28	3,210	7,100	77,700	0,000	2,592	0,000
63	1,02	14,120	40,380	96,680	0,000	1,315	0,000
64	1,18	6,020	7,600	84,000	0,000	7,310	0,000
65	1,50	5,880	7,230	65,530	0,000	4,779	0,000
66	1,28	10,060	20,330	77,630	0,000	17,309	0,000
67	1,25	8,200	11,760	79,670	0,000	26,065	0,000
68	1,17	12,320	32,520	82,950	0,000	0,000	0,000
69	1,00	9,610	44,770	100,000	0,000	0,000	0,000
70	1,14	7,540	14,120	86,750	0,000	6,015	0,000
71	1,18	9,090	22,820	82,260	0,000	0,000	0,000
72	1,20	8,560	15,200	83,190	0,000	26,234	0,000
73	1,07	8,690	31,440	92,980	0,000	3,375	0,000
74	1,07	15,650	49,580	89,780	0,000	0,000	0,000
75	1,38	3,850	7,210	71,960	0,000	0,345	0,000
76	1,21	5,310	3,640	82,090	0,000	22,529	0,000
77	1,18	2,490	1,670	80,040	0,000	0,000	0,000
78	1,00	0,870	1,670	69,280	0,000	0,000	0,000
79	1,08	2,500	16,320	92,120	1,773	0,000	0,000
80	1,00	1,750	9,010	99,230	0,000	0,000	0,000

UNIV	DEA	Observado			Folgas		
		TITULADO GRD	TITULADO PG	IGC	TITULADO GRD	TITULADO PG	IGC
81	1,05	0,860	1,320	94,480	0,837	0,824	0,000
82	1,16	1,030	2,600	83,560	0,034	0,000	0,000
83	1,24	4,640	3,900	79,830	0,000	8,150	0,000
84	1,19	4,500	8,290	83,120	0,000	3,903	0,000
85	1,52	7,020	2,450	56,780	0,000	0,000	0,000
86	2,03	2,290	3,050	48,790	0,000	0,277	0,000
87	1,58	6,110	1,100	52,990	0,000	0,000	0,000
88	1,55	1,970	1,800	64,120	1,001	12,010	0,000
89	1,30	2,470	3,620	76,350	0,000	3,792	0,000
90	1,33	5,760	7,230	73,460	0,000	0,000	0,000
91	1,53	2,540	1,600	64,190	0,000	0,000	0,000
92	1,00	0,570	1,080	91,610	0,000	0,000	0,000
93	1,44	2,390	1,020	59,710	0,000	0,000	0,000
94	1,67	1,600	0,610	56,800	0,000	0,311	0,000
95	1,00	0,400	5,130	91,530	0,000	0,000	0,000
96	1,23	3,890	9,070	81,290	0,255	8,808	0,000
97	1,65	6,660	1,520	58,900	0,000	1,202	0,000
98	1,38	1,560	1,360	72,230	0,000	2,386	0,000
99	1,22	1,670	3,380	79,730	0,000	0,000	0,000
100	1,00	1,590	1,450	99,370	0,000	0,000	0,000

## APÊNDICE F – Frações de produtos por insumo

UNIV	concluintes/docente	titulados/docente	igc/docente	UNIV	concluintes/docente	titulados/docente	igc/docente
1	1,1128	0,2083	0,0017	51	4,2493	0,0444	0,0034
2	1,4863	0,5189	0,0015	52	3,4687	0,1124	0,0012
3	1,4159	0,2563	0,0021	53	1,5101	0,5116	0,0029
4	1,1563	0,1926	0,0015	54	1,4991	0,6809	0,0021
5	1,5478	0,1574	0,0018	55	0,6999	0,2855	0,0009
6	0,8525	0,3023	0,0041	56	1,0936	0,1456	0,0021
7	0,9225	0,6142	0,0033	57	1,3093	0,2948	0,0010
8	1,4120	0,7291	0,0036	58	1,1144	0,3120	0,0014
9	1,2079	0,2490	0,0018	59	0,9968	0,4990	0,0014
10	4,5619	0,2020	0,0019	60	0,9866	0,2581	0,0009
11	1,0985	0,2664	0,0039	61	1,1538	0,3477	0,0019
12	3,4010	0,1054	0,0028	62	0,9609	0,3287	0,0033
13	2,1347	0,3535	0,0034	63	1,3944	0,6161	0,0014
14	1,9910	0,0442	0,0025	64	1,2800	0,2495	0,0026
15	1,4929	0,3343	0,0021	65	1,1453	0,2176	0,0018
16	2,5373	0,2030	0,0089	66	0,8562	0,2673	0,0009
17	3,2537	0,0539	0,0028	67	0,9839	0,2180	0,0014
18	2,4582	0,1145	0,0029	68	1,4451	0,5890	0,0014
19	2,5595	0,4633	0,0025	69	1,1183	0,8047	0,0017
20	3,7232	0,1800	0,0020	70	1,3035	0,3769	0,0021
21	0,8918	0,1308	0,0036	71	1,5153	0,5878	0,0020
22	1,5545	0,3107	0,0023	72	0,8907	0,2443	0,0012
23	1,6206	0,2571	0,0042	73	1,1360	0,6349	0,0017
24	1,2883	0,7855	0,0022	74	1,2348	0,6041	0,0010
25	1,8980	1,1988	0,0010	75	1,0930	0,3159	0,0029
26	1,2227	0,3431	0,0021	76	0,9303	0,0984	0,0021
27	1,4764	0,1302	0,0027	77	1,8329	0,1892	0,0084
28	2,0774	0,1246	0,0019	78	1,0569	0,3130	0,0121
29	2,6007	0,1060	0,0108	79	0,6151	0,6201	0,0032
30	4,6552	0,1322	0,0072	80	0,9242	0,7337	0,0075
31	9,2573	0,0511	0,0008	81	0,5343	0,1268	0,0084
32	4,7868	0,2640	0,0043	82	0,7649	0,2970	0,0089
33	5,5693	0,0591	0,0050	83	1,1723	0,1520	0,0029
34	2,7991	0,0525	0,0067	84	1,1427	0,3254	0,0030
35	5,8439	0,2563	0,0056	85	3,6190	0,1948	0,0042
36	3,0312	0,2305	0,0043	86	0,7867	0,1617	0,0024
37	1,9030	0,1612	0,0048	87	3,5290	0,0985	0,0044
38	6,7123	0,0253	0,0010	88	0,5369	0,0755	0,0025
39	3,3574	0,0089	0,0004	89	0,9058	0,2044	0,0040
40	4,1268	0,2203	0,0016	90	1,4684	0,2850	0,0027
41	3,1402	0,1320	0,0049	91	1,3041	0,1271	0,0047
42	3,3410	0,0922	0,0062	92	0,6840	0,2000	0,0157
43	2,5184	0,2304	0,0033	93	2,1636	0,1424	0,0078
44	1,6305	0,0595	0,0028	94	1,1975	0,0700	0,0061
45	4,2405	0,0696	0,0019	95	0,4075	0,8116	0,0134
46	7,3587	0,0691	0,0016	96	0,8601	0,3101	0,0026
47	6,2334	0,0448	0,0034	97	1,5049	0,0529	0,0019
48	2,1214	0,0765	0,0035	98	0,7962	0,1079	0,0053
49	3,6573	0,1122	0,0047	99	1,0846	0,3384	0,0074
50	4,5224	0,1233	0,0056	100	1,0721	0,1509	0,0096



**APÊNDICE G – Metas proporcionais e globais**

UNIV	DEA	Meta proporcional			Meta Global		
		TITULADO GRD	TITULADO PG	IGC	TITULADO GRD	TITULADO PG	IGC
1	1,37	9,453	11,453	98,366	9,453	23,492	98,366
2	1,08	13,489	30,488	97,859	13,489	32,764	97,859
3	1,40	9,268	10,850	97,874	9,268	10,850	97,874
4	1,56	10,624	11,450	97,843	10,624	17,716	97,843
5	1,40	11,970	7,882	97,454	11,970	11,063	97,454
6	1,25	2,950	6,763	99,588	3,050	9,261	99,588
7	1,06	4,007	17,278	99,375	4,530	17,343	99,375
8	1,03	5,583	18,664	98,035	5,583	18,664	98,035
9	1,21	9,402	12,548	98,942	9,402	25,868	98,942
10	1,24	28,284	8,110	83,737	28,284	8,110	83,737
11	1,34	4,047	6,352	99,173	4,047	6,525	99,173
12	1,35	15,741	3,159	91,652	15,741	3,159	91,652
13	1,19	8,580	9,199	96,093	8,580	9,199	96,093
14	1,67	11,072	1,587	95,207	11,072	1,587	95,207
15	1,18	10,030	14,538	98,294	10,030	17,620	98,294
16	1,22	3,465	1,793	85,156	3,465	1,793	85,156
17	1,49	15,332	1,639	91,397	15,332	1,639	91,397
18	1,39	11,468	3,461	95,007	11,468	3,461	95,007
19	1,10	13,618	15,961	94,501	13,618	15,961	94,501
20	1,29	23,568	7,379	87,423	23,568	7,379	87,423
21	1,41	3,539	3,356	99,123	3,539	8,392	99,123
22	1,52	9,211	11,917	96,824	9,211	11,917	96,824
23	1,14	5,404	5,552	98,188	5,404	5,552	98,188
24	1,00	8,350	32,970	98,410	8,350	32,970	98,410
25	1,00	24,460	100,000	87,660	24,460	100,000	87,660
26	1,18	8,260	14,998	99,073	8,260	22,470	99,073
27	1,48	7,578	4,322	98,124	7,578	4,322	98,124
28	1,42	14,654	5,694	94,742	14,654	5,694	94,742
29	1,00	2,460	0,650	71,190	2,460	0,650	71,190
30	1,00	5,420	1,000	58,710	5,420	1,000	58,710
31	1,00	78,200	2,790	48,930	78,200	2,790	48,930
32	1,21	11,447	4,090	71,777	11,447	4,090	78,755
33	1,12	11,290	0,773	70,616	11,290	1,218	70,616
34	1,34	5,494	0,670	91,401	5,494	1,296	91,401
35	1,00	10,140	2,880	68,050	10,140	2,880	68,050
36	1,54	9,009	4,435	89,335	9,009	4,435	89,335
37	1,36	5,535	3,033	97,689	5,535	3,033	97,689
38	1,00	64,670	1,580	64,630	64,670	1,580	64,630
39	1,00	100,000	1,710	73,570	100,000	1,710	73,570
40	1,25	31,550	10,900	83,450	31,550	10,900	83,450

UNIV	DEA	Meta proporcional			Meta Global		
		TITULADO GRD	TITULADO PG	IGC	TITULADO GRD	TITULADO PG	IGC
41	1,58	8,042	2,196	87,880	8,042	2,196	87,880
42	1,41	6,839	1,227	88,407	6,839	1,236	88,407
43	1,43	10,453	6,192	94,838	10,453	6,192	94,838
44	1,54	8,039	1,894	97,143	8,039	1,894	97,143
45	1,46	26,776	2,847	83,848	26,776	2,847	83,848
46	1,13	31,787	1,932	49,313	31,787	1,932	56,732
47	1,14	15,379	0,718	58,094	15,379	1,252	58,094
48	1,56	8,393	1,966	95,753	8,393	1,966	95,753
49	1,49	9,089	1,803	82,442	9,089	1,803	82,442
50	1,19	8,033	1,416	69,901	8,033	1,416	69,901
51	1,53	15,178	1,025	84,441	15,178	1,441	84,441
52	1,60	32,864	6,896	82,128	32,864	6,896	82,128
53	1,12	7,302	16,027	98,683	7,302	16,027	98,683
54	1,11	9,801	28,816	97,136	9,801	28,816	97,136
55	1,41	10,885	28,750	99,786	10,885	44,215	99,786
56	1,43	7,536	6,507	98,813	7,536	16,183	98,813
57	1,40	17,360	25,312	96,894	17,360	32,401	96,894
58	1,16	11,380	20,613	99,122	11,380	41,951	99,122
59	1,20	9,924	32,160	99,276	9,924	41,037	99,276
60	1,25	15,050	25,488	98,575	15,050	42,227	98,575
61	1,34	8,603	16,804	99,066	8,603	20,249	99,066
62	1,28	4,109	9,088	99,456	4,109	11,680	99,456
63	1,02	14,402	41,188	98,614	14,402	42,503	98,614
64	1,18	7,104	8,968	99,120	7,104	16,278	99,120
65	1,50	8,820	10,845	98,295	8,820	15,624	98,295
66	1,28	12,877	26,022	99,366	12,877	43,332	99,366
67	1,25	10,250	14,700	99,588	10,250	40,765	99,588
68	1,17	14,414	38,048	97,052	14,414	38,048	97,052
69	1,00	9,610	44,770	100,000	9,610	44,770	100,000
70	1,14	8,596	16,097	98,895	8,596	22,112	98,895
71	1,18	10,726	26,928	97,067	10,726	26,928	97,067
72	1,20	10,272	18,240	99,828	10,272	44,474	99,828
73	1,07	9,298	33,641	99,489	9,298	37,016	99,489
74	1,07	16,746	53,051	96,065	16,746	53,051	96,065
75	1,38	5,313	9,950	99,305	5,313	10,295	99,305
76	1,21	6,425	4,404	99,329	6,425	26,934	99,329
77	1,18	2,938	1,971	94,447	2,938	1,971	94,447
78	1,00	0,870	1,670	69,280	0,870	1,670	69,280
79	1,08	2,700	17,626	99,490	4,473	17,626	99,490
80	1,00	1,750	9,010	99,230	1,750	9,010	99,230

UNIV	DEA	Meta proporcional			Meta Global		
		TITULADO GRD	TITULADO PG	IGC	TITULADO GRD	TITULADO PG	IGC
81	1,05	0,903	1,386	99,204	1,740	2,210	99,204
82	1,16	1,195	3,016	96,930	1,229	3,016	96,930
83	1,24	5,754	4,836	98,989	5,754	12,986	98,989
84	1,19	5,355	9,865	98,913	5,355	13,768	98,913
85	1,52	10,670	3,724	86,306	10,670	3,724	86,306
86	2,03	4,649	6,192	99,044	4,649	6,469	99,044
87	1,58	9,654	1,738	83,724	9,654	1,738	83,724
88	1,55	3,054	2,790	99,386	4,054	14,800	99,386
89	1,30	3,211	4,706	99,255	3,211	8,498	99,255
90	1,33	7,661	9,616	97,702	7,661	9,616	97,702
91	1,53	3,886	2,448	98,211	3,886	2,448	98,211
92	1,00	0,570	1,080	91,610	0,570	1,080	91,610
93	1,44	3,442	1,469	85,982	3,442	1,469	85,982
94	1,67	2,672	1,019	94,856	2,672	1,330	94,856
95	1,00	0,400	5,130	91,530	0,400	5,130	91,530
96	1,23	4,785	11,156	99,987	5,039	19,964	99,987
97	1,65	10,989	2,508	97,185	10,989	3,710	97,185
98	1,38	2,153	1,877	99,677	2,153	4,263	99,677
99	1,22	2,037	4,124	97,271	2,037	4,124	97,271
100	1,00	1,590	1,450	99,370	1,590	1,450	99,370