

CODERHOUSE

Education: Higher Education Data 2020

Gustavo Maldonado Saffiotti

2023

Summary

1. Introduction.....	3
2. Data theme	3
3. Goal	3
4. Entity-Relationship Diagram (DER).....	3
5. Table Structure with Definition of Primary and Foreign Keys	4
6. Full Description of Columns by Table.....	5
7. Functional Dashboard Analysis.....	9
7.1. Tab - "Menu"	9
7.2. Tab - "Region"	10
7.3. Tab - "IES"	10
7.4. Tab - "Joiners"	11
8. Measurements Calculated with DAX	12
8.1. Number of Distinct Courses.....	12
8.2. Total number of subscribers.....	12
8.3. Percentage of men enrolled	12
8.4. Percentage of women enrolled	12
8.5. Number of entrants in the age group 0-29 years	12
8.6. Number of entrants in the 30-49 age group	13
8.7. Number of entrants over 50 years old.....	13
8.8. Percentage of ENEM enrollees.....	13
9. Columns Calculated with DAX	13
9.1. "Municipality and State".....	13
9.2. "Valid Courses"	14
10. Reach	14
12. Technological tools implemented.....	14
13. Future guidelines	14
14. Data Source.....	15

1. Introduction

This project was carried out throughout the Data Analytics course, involving multiple steps, from data collection and preparation to modeling and, finally, the presentation of the collected information. The choice of the theme was focused on Education, specifically in the context of higher education in Brazil, with data coming from the National Institute of Educational Studies and Research Anísio Teixeira (Inep), an official government source.

The preparation and modeling of the data were carried out with the aid of Microsoft Excel, where the initial cleaning, standardization and filtering of the data was carried out. Then, an Entity-Relationship Diagram (DER) was developed to structure the database, identifying the entities and their relationships, facilitating analytical modeling and ensuring data consistency.

Finally, the Power BI tool was used to build and enhance dynamic visualizations, allowing for interactive data analysis. With the creation of a Mockup and the application of columns and calculated measures, the main indicators on Brazilian higher education in 2020 were presented in a clear and accessible way.

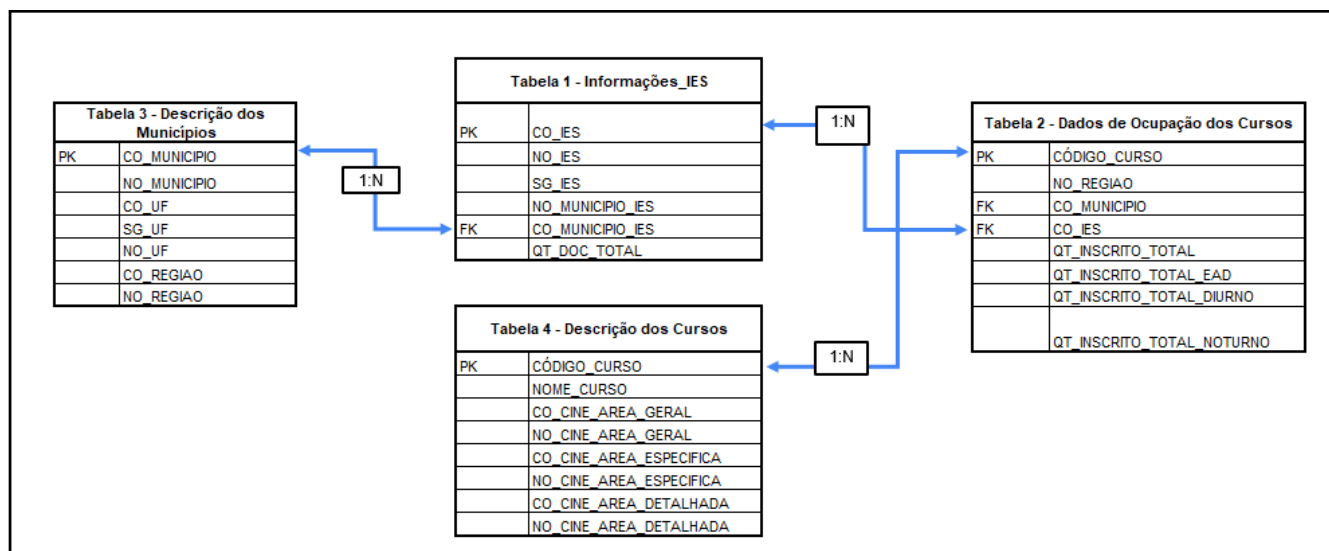
2. Data theme

The theme of this project addresses public higher education in Brazil in 2020, focusing on data related to institutions and courses offered. The dataset provides detailed information on freshmen in the period, including characteristics of the institutions and occupations of the courses, enabling a comprehensive analysis of the profile and distribution of students in Brazilian public higher education. ^[1]

3. Goal

The objective of this project is to analyze data from Brazilian higher education in 2020, identifying patterns and trends related to the preferences of undergraduate courses by region, age group and gender of students. In addition, it seeks to present the number of entrants and the number of educational institutions by region, among other relevant information, offering a detailed view of public higher education in the country.

4. Entity-Relationship Diagram (DER)



Source: The author.

Note: In tables 1 and 2, only the columns that are essential for relationships and those that are most relevant for analysis have been included, due to the large number of columns available. This was done to simplify visualization and focus on the information that is most important to the project.

5. Table Structure with Definition of Primary and Foreign Keys

Tabela 1 - Informações_IES		
Tipo de Chave	Campo	Tipo de campo
PK	CO_IES	int
	NO_IES	nvarchar(100)
	SG_IES	nvarchar(100)
	NO_MUNICIPIO_IES	nvarchar(100)
FK	CO_MUNICIPIO_IES	int
	QT_DOC_TOTAL	int

Tabela 3 - Descrição dos Municípios		
Tipo de Chave	Campo	Tipo de campo
PK	CO_MUNICIPIO	int
	NO_MUNICIPIO	nvarchar(100)
	CO_UF	int
	SG_UF	nvarchar(100)
	NO_UF	nvarchar(100)
	CO_REGIAO	int
	NO_REGIAO	nvarchar(100)

Tabela 2 - Dados de Ocupação dos Cursos		
Tipo de Chave	Campo	Tipo de campo
PK	CÓDIGO_CURSO	int
	NO_REGIAO	nvarchar(100)
FK	CO_MUNICIPIO	int
FK	CO_IES	int
	QT_INSCRITO_TOTAL	int
	QT_INSCRITO_TOTAL_EAD	int
	QT_INSCRITO_TOTAL_DIURNO	int
	QT_INSCRITO_TOTAL_NOTURNO	int

Tabela 4 - Descrição dos Cursos		
Tipo de Chave	Campo	Tipo de campo
PK	CÓDIGO_CURSO	int
	NOME_CURSO	nvarchar(100)
	CO_CINE_AREA_GERAL	int
	NO_CINE_AREA_GERAL	nvarchar(100)
	CO_CINE_AREA_ESPECIFICA	int
	NO_CINE_AREA_ESPECIFICA	nvarchar(100)
	CO_CINE_AREA_DETALHADA	int
	NO_CINE_AREA_DETALHADA	nvarchar(100)

Source: The author.

6. Full Description of Columns by Table

Table 1 – Informações_IES

Nome_coluna	Descrição	Tipo
NU_ANO_CENSO	Ano de referência do Censo da Educação Superior	int
NO_REGIAO_IES	Nome da região geográfica da sede administrativa ou reitoria da IES	nvarchar
CO_REGIAO_IES	Código da região geográfica	int
NO_UF_IES	Nome da Unidade da Federação da sede administrativa ou reitoria da IES	nvarchar
SG_UF_IES	Sigla da Unidade da Federação da sede administrativa ou reitoria da IES	nvarchar
CO_UF_IES	Código da UF da sede administrativa ou reitoria da IES	int
NO_MUNICIPIO_IES	Nome do Município da sede administrativa ou reitoria da IES	nvarchar
CO_MUNICIPIO_IES	Código do Município da sede administrativa ou reitoria da IES	int
IN_CAPITAL_IES	Sede administrativa ou reitoria da IES está localizada em uma capital da unidade da federação	int
NO_MESORREGIAO_IES	Nome da Mesorregião da sede administrativa ou reitoria da IES	nvarchar
CO_MESORREGIAO_IES	Código da Mesorregião da sede administrativa ou reitoria da IES	int
NO_MICRORREGIAO_IES	Nome da Microrregião da sede administrativa ou reitoria da IES	nvarchar
CO_MICRORREGIAO_IES	Código da Microrregião da sede administrativa ou reitoria da IES	int
TP_ORGANIZACAO_ACADEMICA	Tipo de Organização Acadêmica da IES	int
TP_CATEGORIA_ADMINISTRATIVA	Tipo de Categoria Administrativa da IES	int
NO_MANTENEDORA	Nome da mantenedora da IES	nvarchar
CO_MANTENEDORA	Código único de identificação da mantenedora da IES	int
CO_IES	Código único de identificação da IES	int
NO_IES	Nome da IES	nvarchar
SG_IES	Sigla da IES	nvarchar
DS_ENDERECO_IES	Endereço da sede administrativa/reitoria da IES	nvarchar
DS_NUMERO_ENDERECO_IES	Número do endereço	nvarchar
DS_COMPLEMENTO_ENDERECO_IES	Complemento	nvarchar
NO_BAIRRO_IES	Bairro	nvarchar
NU_CEP_IES	CEP	int
QT_TEC_TOTAL	Quantidade de funcionários técnico-administrativos	int
QT_TEC_FUNDAMENTAL_INCOMP_FEM	Quantidade de funcionários técnico-administrativos do sexo feminino com ensino fundamental incompleto	int
QT_TEC_FUNDAMENTAL_INCOMP_MASC	Quantidade de funcionários técnico-administrativos do sexo masculino com ensino fundamental incompleto	int
QT_TEC_FUNDAMENTAL_COMP_FEM	Quantidade de funcionários técnico-administrativos do sexo feminino com ensino fundamental completo	int
QT_TEC_FUNDAMENTAL_COMP_MASC	Quantidade de funcionários técnico-administrativos do sexo masculino com ensino fundamental completo	int
QT_TEC_MEDIO_FEM	Quantidade de funcionários técnico-administrativos do sexo feminino com ensino médio	int
QT_TEC_MEDIO_MASC	Quantidade de funcionários técnico-administrativos do sexo masculino com ensino médio	int
QT_TEC_SUPERIOR_FEM	Quantidade de funcionários técnico-administrativos do sexo feminino com nível superior	int
QT_TEC_SUPERIOR_MASC	Quantidade de funcionários técnico-administrativos do sexo masculino com nível superior	int
QT_TEC_ESPECIALIZACAO_FEM	Quantidade de funcionários técnico-administrativos do sexo feminino com especialização	int
QT_TEC_ESPECIALIZACAO_MASC	Quantidade de funcionários técnico-administrativos do sexo masculino com especialização	int
QT_TEC_MESTRADO_FEM	Quantidade de funcionários técnico-administrativos do sexo feminino com mestrado	int
QT_TEC_MESTRADO_MASC	Quantidade de funcionários técnico-administrativos do sexo masculino com mestrado	int
QT_TEC_DOUTORADO_FEM	Quantidade de funcionários técnico-administrativos do sexo feminino com doutorado	int
QT_TEC_DOUTORADO_MASC	Quantidade de funcionários técnico-administrativos do sexo masculino com doutorado	int
IN_ACESSO_PORTAL_CAPES	Informa se as bibliotecas da IES têm acesso ao portal Capes de periódicos	int
IN_ACESSO_OUTRAS_BASES	Informa se as bibliotecas da IES têm acesso a outras bases de dados licenciadas ou compradas	int
IN_ASSINA_OUTRA_BASE	Informa se as bibliotecas da IES assinam outras bases de dados licenciadas ou compradas	int
IN_REPOSITORIO_INSTITUCIONAL	Informa se a IES possui base de dados online que reúne de maneira organizada a produção científica da instituição	int
IN_BUSCA_INTEGRADA	Informa se a IES oferece ferramenta eletrônica que possibilita pesquisar simultaneamente diversas fontes de informação	int
IN_SERVICO_INTERNET	Informa se as bibliotecas da IES oferecem serviços pela internet	int
IN_PARTICIPA_REDE_SOCIAL	Informa se as bibliotecas participam de Redes Sociais	int
IN_CATALOGO_ONLINE	Informa se a IES possui ferramenta de recuperação de informação que permite ao usuário consultar, de forma integrada, os diversos tipos de documentos disponíveis	int
QT_PERIODICO_ELETRONICO	Quantidade de títulos de periódicos eletrônicos adquiridos pelas bibliotecas por meio de compra, doação ou permuta	int
QT_LIVRO_ELETRONICO	Quantidade de títulos de livros eletrônicos disponibilizados pelas bibliotecas convertidos ao formato digital ou em PDF	int
QT_DOC_TOTAL	Quantidade total de docentes (em exercício e afastados)	int
QT_DOC_EXE	Quantidade total de docentes em exercício	int
QT_DOC_EX_FEMI	Quantidade de docentes em exercício do sexo feminino	int
QT_DOC_EX_MASC	Quantidade de docentes em exercício do sexo masculino	int
QT_DOC_EX_SEM_GRAD	Quantidade de docentes em exercício sem curso de graduação	int
QT_DOC_EX_GRAD	Quantidade de docentes em exercício com curso de graduação	int
QT_DOC_EX_ESP	Quantidade de docentes em exercício com especialização	int
QT_DOC_EX_MEST	Quantidade de docentes em exercício com mestrado	int
QT_DOC_EX_DOUT	Quantidade de docentes em exercício com doutorado	int
QT_DOC_EX_INT	Quantidade de docentes em exercício em tempo integral	int
QT_DOC_EX_INT_DE	Quantidade de docentes em exercício em tempo integral com dedicação exclusiva	int
QT_DOC_EX_INT_SEM_DE	Quantidade de docentes em exercício em tempo integral sem dedicação exclusiva	int
QT_DOC_EX_PARC	Quantidade de docentes em exercício em tempo parcial	int
QT_DOC_EX_HOR	Quantidade de docentes em exercício horista	int
QT_DOC_EX_0_29	Quantidade de docentes em exercício - até 29 anos	int
QT_DOC_EX_30_34	Quantidade de docentes em exercício - de 30 a 34 anos	int
QT_DOC_EX_35_39	Quantidade de docentes em exercício - de 35 a 39 anos	int
QT_DOC_EX_40_44	Quantidade de docentes em exercício - de 40 a 44 anos	int
QT_DOC_EX_45_49	Quantidade de docentes em exercício - de 45 a 49 anos	int
QT_DOC_EX_50_54	Quantidade de docentes em exercício - de 50 a 54 anos	int
QT_DOC_EX_55_59	Quantidade de docentes em exercício - de 55 a 59 anos	int
QT_DOC_EX_60 MAIS	Quantidade de docentes em exercício - de 60 anos ou mais	int
QT_DOC_EX_BRANCA	Quantidade de docentes em exercício - Cor/Raça branca	int
QT_DOC_EX_PRETA	Quantidade de docentes em exercício - Cor/Raça preta	int
QT_DOC_EX_PARDA	Quantidade de docentes em exercício - Cor/Raça parda	int
QT_DOC_EX_AMARELA	Quantidade de docentes em exercício - Cor/Raça amarela	int
QT_DOC_EX_INDIGENA	Quantidade de docentes em exercício - Cor/Raça indígena	int
QT_DOC_EX_COR_ND	Quantidade de docentes em exercício - Cor/Raça não dispõe da informação ou não declarada	int
QT_DOC_EX_BRA	Quantidade de docentes em exercício - nacionalidade brasileira ou brasileira - nascido no exterior ou naturalizado	int
QT_DOC_EX_EST	Quantidade de docentes em exercício - nacionalidade estrangeira	int
QT_DOC_EX_COM_DEFICIENCIA	Quantidade de docentes em exercício - com deficiência	int

Source: The author.

Table 2 – Course Occupancy Data

Nome_coluna	Descrição	Tipo
NU_ANO_CENSO	Ano de referência do Censo da Educação Superior	int
NO_REGIAO	Nome da região geográfica do local de oferta do curso	nvarchar
CO_REGIAO	Código da região geográfica	int
NO_UF	Nome da Unidade da Federação do local de oferta do curso	nvarchar
SG_UF	Sigla da Unidade da Federação do local de oferta do curso	nvarchar
CO_UF	Código do IBGE da Unidade da Federação do local de oferta do curso	int
NO_MUNICIPIO	Nome do Município do local de oferta do curso	nvarchar
CO_MUNICIPIO	Código do IBGE do Município do local de oferta do curso	int
IN_CAPITAL	Informa se o local de oferta do curso está localizado em capital de Estado	int
TP_ORGANIZACAO_ACADEMICA	Tipo da Organização Acadêmica da IES	int
TP_CATEGORIA_ADMINISTRATIVA	Tipo da Categoria Administrativa da IES	int
TP_REDE	Rede de Ensino	int
CO_IES	Código da Instituição	int
NO_CINE_ROTULO	Nome do curso, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
CODIGO_CURSO	Código de identificação do curso, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
CO_CINE_AREA_GERAL	Código de identificação da área geral, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
NO_CINE_AREA_GERAL	Nome da área geral, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
CO_CINE_AREA_ESPECIFICA	Código de identificação da área específica, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
NO_CINE_AREA_ESPECIFICA	Nome da área específica, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
CO_CINE_AREA_DETALHADA	Código de identificação da área detalhada, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
NO_CINE_AREA_DETALHADA	Nome da área detalhada, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesc	nvarchar
TP_GRAU_ACADEMICO	Tipo do grau acadêmico conferido ao aluno pela conclusão dos requisitos exigidos pelo curso	int
TP_MODALIDADE_ENSINO	Tipo da modalidade de ensino do curso	int
TP_NIVEL_ACADEMICO	Tipo do nível acadêmico do curso	int
QT_CURSO	Número de cursos	int
QT_VG_TOTAL	Quantidade total de vagas oferecidas nos cursos	int
QT_VG_TOTAL_DIURNO	Quantidade total de vagas oferecidas nos cursos no turno diurno	int
QT_VG_TOTAL_NOTURNO	Quantidade total de vagas oferecidas nos cursos no turno noturno	int
QT_VG_TOTAL_EAD	Quantidade total de vagas oferecidas nos cursos EAD	int
QT_VG_NOVA	Quantidade de vagas novas oferecidas nos cursos	int
QT_VG_PROC_SELETIVO	Quantidade de vagas nos processos seletivos oferecidos nos cursos	int
QT_VG_REMANESC	Quantidade de vagas remanescentes oferecidas nos cursos	int
QT_VG_PROG_ESPECIAL	Quantidade de vagas oferecidas para programas especiais	int
QT_INSCRITO_TOTAL	Quantidade total de inscritos nos cursos	int
QT_INSCRITO_TOTAL_DIURNO	Quantidade total de inscritos nos cursos no turno diurno	int
QT_INSCRITO_TOTAL_NOTURNO	Quantidade total de inscritos nos cursos no turno noturno	int
QT_INSCRITO_TOTAL_EAD	Quantidade total de inscritos nos cursos EAD	int
QT_INSC_VG_NOVA	Quantidade de inscritos no processo seletivo de vagas novas	int
QT_INSC_PROC_SELETIVO	Quantidade de inscritos nos processos seletivos oferecidos nos cursos	int
QT_INSC_VG_REMANESC	Quantidade de inscritos no processo seletivo de vagas remanescentes	int
QT_INSC_VG_PROG_ESPECIAL	Quantidade de inscritos no processo seletivo de vagas oferecidas para programas especiais	int
QT_ING	Quantidade de ingressantes nos cursos	int
QT_ING_FEM	Quantidade de ingressantes do sexo feminino nos cursos	int
QT_ING_MASC	Quantidade de ingressantes do sexo masculino nos cursos	int
QT_ING_DIURNO	Quantidade de ingressantes nos cursos no turno diurno	int
QT_ING_NOTURNO	Quantidade de ingressantes nos cursos no turno noturno	int
QT_ING_VG_NOVA	Quantidade de ingressantes em vagas novas oferecidas nos cursos	int
QT_ING_VESTIBULAR	Quantidade de ingressantes nos cursos por vestibular	int
QT_ING_ENEM	Quantidade de ingressantes nos cursos pelo ENEM	int
QT_ING_AVALIACAO_SERIADA	Quantidade de ingressantes nos cursos por meio da avaliação seriada	int
QT_ING_SELECAO_SIMPLIFICA	Quantidade de ingressantes nos cursos por meio de seleção simplificada	int
QT_ING_EGR	Quantidade de alunos que se formaram em um curso de Bacharelado Interdisciplinar ou Licenciatura Interdisciplinar	int
QT_ING_OUTRO_TIPO_SELECAO	Quantidade de ingressantes em outro tipo de seleção para os cursos	int
QT_ING_PROC_SELETIVO	Quantidade de ingressantes nos processos seletivos oferecidos nos cursos	int
QT_ING_VG_REMANESC	Quantidade de ingressantes nos cursos por meio de vagas remanescentes	int
QT_ING_VG_PROG_ESPECIAL	Quantidade de ingressantes nos cursos por meio de vagas para programas especiais	int
QT_ING_OUTRA_FORMA	Quantidade de ingressantes nos cursos por outras formas de ingresso	int
QT_ING_0_17	Quantidade de ingressantes nos cursos - Até 17 anos de idade	int
QT_ING_18_24	Quantidade de ingressantes nos cursos - De 18 a 24 anos de idade	int
QT_ING_25_29	Quantidade de ingressantes nos cursos - De 25 a 29 anos de idade	int
QT_ING_30_34	Quantidade de ingressantes nos cursos - De 30 a 34 anos de idade	int
QT_ING_35_39	Quantidade de ingressantes nos cursos - De 35 a 39 anos de idade	int
QT_ING_40_49	Quantidade de ingressantes nos cursos - De 40 a 49 anos de idade	int
QT_ING_50_59	Quantidade de ingressantes nos cursos - De 50 a 59 anos de idade	int
QT_ING_60 MAIS	Quantidade de ingressantes nos cursos - De 60 ou mais anos de idade	int
QT_ING_BRANCA	Quantidade de ingressantes nos cursos - Cor/Raça Branca	int
QT_ING_PRETA	Quantidade de ingressantes nos cursos - Cor/Raça Preta	int
QT_ING_PARDA	Quantidade de ingressantes nos cursos - Cor/Raça Parda	int
QT_ING_AMARELA	Quantidade de ingressantes nos cursos - Cor/Raça Amarela	int
QT_ING_INDIGENA	Quantidade de ingressantes nos cursos - Cor/Raça Indígena	int
QT_ING_CORND	Quantidade de ingressantes nos cursos - Cor/Raça Não Declarada	int
QT_MAT	Quantidade de matrículas	int
QT_MAT_FEM	Quantidade de matrículas do sexo feminino	int
QT_MAT_MASC	Quantidade de matrículas do sexo masculino	int
QT_MAT_DIURNO	Quantidade de matrículas - turno diurno	int
QT_MAT_NOTURNO	Quantidade de matrículas - turno noturno	int
QT_MAT_0_17	Quantidade de matrículas - Até 17 anos de idade	int
QT_MAT_18_24	Quantidade de matrículas - De 18 a 24 anos de idade	int
QT_MAT_25_29	Quantidade de matrículas - De 25 a 29 anos de idade	int

QT_MAT_30_34	Quantidade de matrículas - De 30 a 34 anos de idade	int
QT_MAT_35_39	Quantidade de matrículas - De 35 a 39 anos de idade	int
QT_MAT_40_49	Quantidade de matrículas - De 40 a 49 anos de idade	int
QT_MAT_50_59	Quantidade de matrículas - De 50 a 59 anos de idade	int
QT_MAT_60 MAIS	Quantidade de matrículas - De 60 ou mais anos de idade	int
QT_MAT_BRANCA	Quantidade de matrículas - Cor/Raça Branca	int
QT_MAT_PRETA	Quantidade de matrículas - Cor/Raça Preta	int
QT_MAT_PARDA	Quantidade de matrículas - Cor/Raça Parda	int
QT_MAT_AMARELA	Quantidade de matrículas - Cor/Raça Amarela	int
QT_MAT_INDIGENA	Quantidade de matrículas - Cor/Raça Indígena	int
QT_MAT_CORND	Quantidade de matrículas - Cor/Raça Não Declarada	int
QT_CONC	Quantidade de concluintes	int
QT_CONC_FEM	Quantidade de concluintes do sexo feminino	int
QT_CONC_MASC	Quantidade de concluintes do sexo masculino	int
QT_CONC_DIURNO	Quantidade de concluintes - turno diurno	int
QT_CONC_NOTURNO	Quantidade de concluintes - turno noturno	int
QT_CONC_0_17	Quantidade de concluintes - Até 17 anos de idade	int
QT_CONC_18_24	Quantidade de concluintes - De 18 a 24 anos de idade	int
QT_CONC_25_29	Quantidade de concluintes - De 25 a 29 anos de idade	int
QT_CONC_30_34	Quantidade de concluintes - De 30 a 34 anos de idade	int
QT_CONC_35_39	Quantidade de concluintes - De 35 a 39 anos de idade	int
QT_CONC_40_49	Quantidade de concluintes - De 40 a 49 anos de idade	int
QT_CONC_50_59	Quantidade de concluintes - De 50 a 59 anos de idade	int
QT_CONC_60 MAIS	Quantidade de concluintes - De 60 ou mais anos de idade	int
QT_CONC_BRANCA	Quantidade de concluintes - Cor/Raça Branca	int
QT_CONC_PRETA	Quantidade de concluintes - Cor/Raça Preta	int
QT_CONC_PARDA	Quantidade de concluintes - Cor/Raça Parda	int
QT_CONC_AMARELA	Quantidade de concluintes - Cor/Raça Amarela	int
QT_CONC_INDIGENA	Quantidade de concluintes - Cor/Raça Indígena	int
QT_CONC_CORND	Quantidade de concluintes - Cor/Raça Não Declarada	int
QT_ING_NACBRAS	Quantidade de ingressantes nos cursos - nacionalidade brasileira ou brasileira - nascido no exterior ou n	int
QT_ING_NACESTRANG	Quantidade de ingressantes nos cursos - nacionalidade estrangeira	int
QT_MAT_NACBRAS	Quantidade de matrículas - nacionalidade brasileira ou brasileira - nascido no exterior ou naturalizado	int
QT_MAT_NACESTRANG	Quantidade de matrículas - nacionalidade estrangeira	int
QT_CONC_NACBRAS	Quantidade de concluintes - nacionalidade brasileira ou brasileira - nascido no exterior ou naturalizado	int
QT_CONC_NACESTRANG	Quantidade de concluintes - nacionalidade estrangeira	int
QT_ALUNO_DEFICIENTE	Quantidade de alunos com deficiência, transtorno global do desenvolvimento ou altas habilidades/supe	int
QT_ING_DEFICIENTE	Quantidade de ingressantes nos cursos - alunos com deficiência, transtorno global do desenvolvimento	int
QT_MAT_DEFICIENTE	Quantidade de matrículas - alunos com deficiência, transtorno global do desenvolvimento ou altas habil	int
QT_CONC_DEFICIENTE	Quantidade de concluintes - alunos com deficiência, transtorno global do desenvolvimento ou altas habil	int
QT_ING_FINANC	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil	int
QT_ING_FINANC_REEMB	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil reembolsável	int
QT_ING_FIES	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil reembolsável - F	int
QT_ING_RPFIES	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil reembolsável - ac	int
QT_ING_FINANC_REEMB_OUTROS	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil reembolsável - pe	int
QT_ING_FINANC_NREEMB	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil não reembolsáve	int
QT_ING_PROUNII	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil não reembolsáve	int
QT_ING_PROUNIP	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil não reembolsáve	int
QT_ING_NRPFI	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil não reembolsáve	int
QT_ING_FINANC_NREEMB_OUTROS	Quantidade de ingressantes nos cursos - alunos que utilizam financiamento estudantil não reembolsáve	int
QT_MAT_FINANC	Quantidade de matrículas - alunos que utilizam financiamento estudantil	int
QT_MAT_FINANC_REEMB	Quantidade de matrículas - alunos que utilizam financiamento estudantil reembolsável	int
QT_MAT_FIES	Quantidade de matrículas - alunos que utilizam financiamento estudantil reembolsável - FIES	int
QT_MAT_RPFIES	Quantidade de matrículas - alunos que utilizam financiamento estudantil reembolsável - administrado p	int
QT_MAT_FINANC_REEMB_OUTROS	Quantidade de matrículas - alunos que utilizam financiamento estudantil reembolsável - pelo governo e	int
QT_MAT_FINANC_NREEMB	Quantidade de matrículas - alunos que utilizam financiamento estudantil não reembolsável	int
QT_MAT_PROUNII	Quantidade de matrículas - alunos que utilizam financiamento estudantil não reembolsável - PROUNI In	int
QT_MAT_PROUNIP	Quantidade de matrículas - alunos que utilizam financiamento estudantil não reembolsável - PROUNI P	int
QT_MAT_NRPFI	Quantidade de matrículas - alunos que utilizam financiamento estudantil não reembolsável - administra	int
QT_MAT_FINANC_NREEMB_OUTROS	Quantidade de matrículas - alunos que utilizam financiamento estudantil não reembolsável - pelo gover	int
QT_CONC_FINANC	Quantidade de concluintes - alunos que utilizam financiamento estudantil	int
QT_CONC_FINANC_REEMB	Quantidade de concluintes - alunos que utilizam financiamento estudantil reembolsável	int
QT_CONC_FIES	Quantidade de concluintes - alunos que utilizam financiamento estudantil reembolsável - FIES	int
QT_CONC_RPFIES	Quantidade de concluintes - alunos que utilizam financiamento estudantil reembolsável - administrado	int
QT_CONC_FINANC_REEMB_OUTROS	Quantidade de concluintes - alunos que utilizam financiamento estudantil reembolsável - pelo governo	int
QT_CONC_FINANC_NREEMB	Quantidade de concluintes - alunos que utilizam financiamento estudantil não reembolsável	int
QT_CONC_PROUNII	Quantidade de concluintes - alunos que utilizam financiamento estudantil não reembolsável - PROUNI I	int
QT_CONC_PROUNIP	Quantidade de concluintes - alunos que utilizam financiamento estudantil não reembolsável - PROUNI f	int
QT_CONC_NRPFI	Quantidade de concluintes - alunos que utilizam financiamento estudantil não reembolsável - administ	int
QT_CONC_FINANC_NREEMB_OUTROS	Quantidade de concluintes - alunos que utilizam financiamento estudantil não reembolsável - pelo gover	int
QT_ING_RESERVA_VAGA	Quantidade de ingressantes nos cursos - alunos que participam do programa de reserva de vagas	int
QT_ING_RVREDEPUBLICA	Quantidade de ingressantes nos cursos - alunos que participam do programa de reserva de vagas - escola	int
QT_ING_RVETNICO	Quantidade de ingressantes nos cursos - alunos que participam do programa de reserva de vagas - cunho	int
QT_ING_RVPDEF	Quantidade de ingressantes nos cursos - alunos que participam do programa de reserva de vagas - pesso	int
QT_ING_RVSOCIAL_RF	Quantidade de ingressantes nos cursos - alunos que participam do programa de reserva de vagas - cunho	int
QT_ING_RVOUTROS	Quantidade de ingressantes nos cursos - alunos que participam do programa de reserva de vagas - outros	int
QT_MAT_RESERVA_VAGA	Quantidade de matrículas - alunos que participam do programa de reserva de vagas	int
QT_MAT_RVREDEPUBLICA	Quantidade de matrículas - alunos que participam do programa de reserva de vagas - escola pública	int
QT_MAT_RVETNICO	Quantidade de matrículas - alunos que participam do programa de reserva de vagas - cunho étnico	int
QT_MAT_RVPDEF	Quantidade de matrículas - alunos que participam do programa de reserva de vagas - pessoas com deficiê	int

QT_MAT_RVSOCIAL_RF	Quantidade de matrículas - alunos que participam do programa de reserva de vagas - cunho social/renda	int
QT_MAT_RVOUTROS	Quantidade de matrículas - alunos que participam do programa de reserva de vagas - outros tipos	int
QT_CONC_RESERVA_VAGA	Quantidade de concluintes - alunos que participam do programa de reserva de vagas	int
QT_CONC_RVREDEPUBLICA	Quantidade de concluintes - alunos que participam do programa de reserva de vagas - escola pública	int
QT_CONC_RVETNICO	Quantidade de concluintes - alunos que participam do programa de reserva de vagas - cunho étnico	int
QT_CONC_RVPDEF	Quantidade de concluintes - alunos que participam do programa de reserva de vagas - pessoas com deficiência	int
QT_CONC_RVSOCIAL_RF	Quantidade de concluintes - alunos que participam do programa de reserva de vagas - cunho social/renda	int
QT_CONC_RVOUTROS	Quantidade de concluintes - alunos que participam do programa de reserva de vagas - outros tipos	int
QT_SIT_TRANCADA	Quantidade de alunos com situação de vínculo no curso - matrícula trancada	int
QT_SIT_DESVINCULADO	Quantidade de alunos com situação de vínculo no curso - aluno desvinculado do curso	int
QT_SIT_TRANSFERIDO	Quantidade de alunos com situação de vínculo no curso - aluno transferido para outro curso da mesma IE	int
QT_SIT_FALECIDO	Quantidade de alunos com situação de vínculo no curso - aluno falecido	int
QT_ING_PROCESCPUBLICA	Quantidade de ingressantes que terminaram o ensino médio em escolas públicas	int
QT_ING_PROCESCPRIVADA	Quantidade de ingressantes que terminaram o ensino médio em escolas privadas	int
QT_ING_PROCNAOINFORMADA	Quantidade de ingressantes que não informaram o tipo de escola que terminaram o ensino médio	int
QT_MAT_PROCESCPUBLICA	Quantidade de matrículas de alunos que terminaram o ensino médio em escolas públicas	int
QT_MAT_PROCESCPRIVADA	Quantidade de matrículas de alunos que terminaram o ensino médio em escolas privadas	int
QT_MAT_PROCNAOINFORMADA	Quantidade de matrículas de alunos que não informaram o tipo de escola que terminaram o ensino médio	int
QT_CONC_PROCESCPUBLICA	Quantidade de concluintes que terminaram o ensino médio em escolas públicas	int
QT_CONC_PROCESCPRIVADA	Quantidade de concluintes que terminaram o ensino médio em escolas privadas	int
QT_CONC_PROCNAOINFORMADA	Quantidade de concluintes que não informaram o tipo de escola que terminaram o ensino médio	int
QT_PARFOR	Quantidade de alunos que participam do programa especial para a formação de professores em exercício	int
QT_ING_PARFOR	Quantidade de ingressantes que participam do programa especial para a formação de professores em exercício	int
QT_MAT_PARFOR	Quantidade de matrículas de alunos que participam do programa especial para a formação de professores em exercício	int
QT_CONC_PARFOR	Quantidade de concluintes que participam do programa especial para a formação de professores em exercício	int
QT_APOIO_SOCIAL	Quantidade de alunos que recebem algum tipo de apoio social	int
QT_ING_APOIO_SOCIAL	Quantidade de ingressantes que recebem algum tipo de apoio social	int
QT_MAT_APOIO_SOCIAL	Quantidade de matrículas de alunos que recebem algum tipo de apoio social	int
QT_CONC_APOIO_SOCIAL	Quantidade de concluintes que recebem algum tipo de apoio social	int
QT_ATIV_EXTRACURRICULAR	Quantidade de alunos que participam de algum tipo de atividade extracurricular	int
QT_ING_ATIV_EXTRACURRICULAR	Quantidade de ingressantes que participam de algum tipo de atividade extracurricular	int
QT_MAT_ATIV_EXTRACURRICULAR	Quantidade de matrículas de alunos que participam de algum tipo de atividade extracurricular	int
QT_CONC_ATIV_EXTRACURRICULAR	Quantidade de concluintes que participam de algum tipo de atividade extracurricular	int
QT_MOB_ACADEMICA	Quantidade de alunos de curso de graduação que se vincularam temporariamente a outra instituição, se	int
QT_ING_MOB_ACADEMICA	Quantidade de alunos ingressantes em curso de graduação que se vincularam temporariamente a outra instituição	int
QT_MAT_MOB_ACADEMICA	Quantidade de alunos que estão regularmente matriculados em cursos de graduação, que se vincularam	int
QT_CONC_MOB_ACADEMICA	Quantidade de concluintes em cursos de graduação que se vincularam temporariamente a outra instituição	int

Source: The author.

Table 3 – Description of Municipalities

Nome_coluna	Descrição	Tipo
NO_REGIAO	Nome da região geográfica do local de oferta do curso	nvarchar
CO_REGIAO	Código da região geográfica	int
NO_UF	Nome da Unidade da Federação do local de oferta do curso	nvarchar
SG_UF	Sigla da Unidade da Federação do local de oferta do curso	nvarchar
CO_UF	Código do IBGE da Unidade da Federação do local de oferta do curso	int
NO_MUNICIPIO	Nome do Município do local de oferta do curso	nvarchar
CO_MUNICIPIO	Código do IBGE do Município do local de oferta do curso	int

Source: The author.

Table 4 – Description of Courses

Nome_coluna	Descrição	Tipo
NOME_CURSO	Nome do curso, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	nvarchar
CÓDIGO_CURSO	Código de identificação do curso, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	int
CO_CINE_AREA_GERAL	Código de identificação da área geral, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	int
NO_CINE_AREA_GERAL	Nome da área geral, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	nvarchar
CO_CINE_AREA_ESPECIFICA	Código de identificação da área específica, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	int
NO_CINE_AREA_ESPECIFICA	Nome da área específica, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	nvarchar
CO_CINE_AREA_DETALHADA	Código de identificação da área detalhada, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	int
NO_CINE_AREA_DETALHADA	Nome da área detalhada, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco	nvarchar

Source: The author.

7. Functional Dashboard Analysis

7.1. Tab - "Menu"

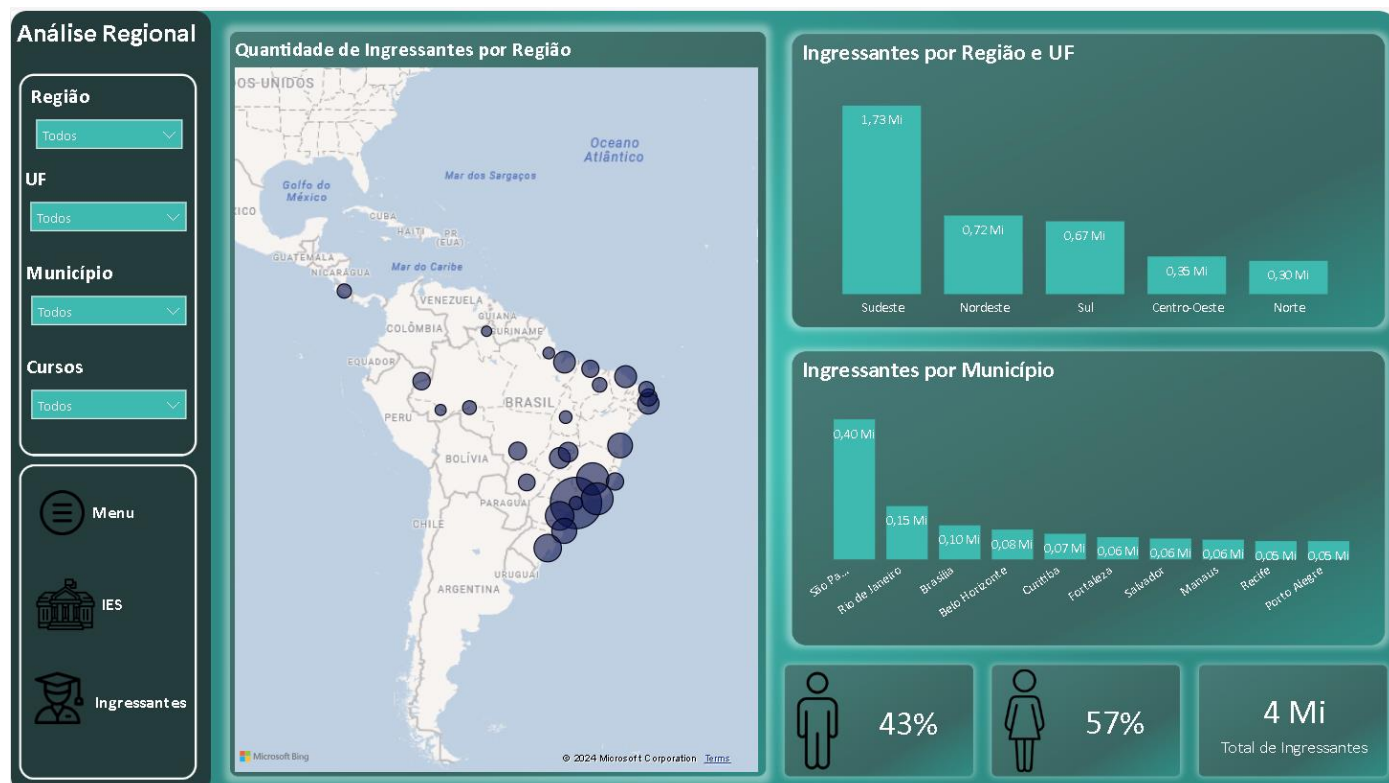
In the first tab, there is the user menu, which features three icons: Region, IES (Higher Education Institutions) and freshmen. These icons allow the user to select the desired analysis, providing intuitive navigation and making it easier to view relevant information according to the focus of interest.



Source: The author.

7.2. Tab - "Region"

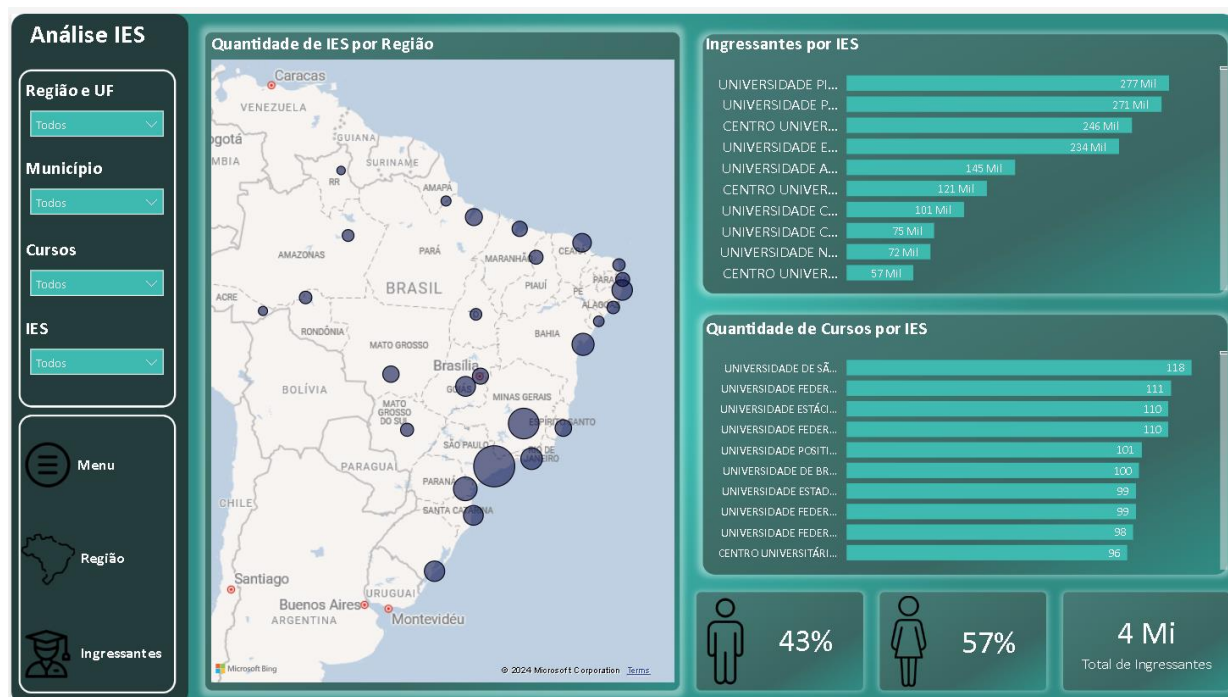
In the second tab, three graphs are presented: the first map graph shows the distribution of freshmen by region, the second shows the number of freshmen by region and state, and the third highlights the 10 municipalities with the highest number of freshmen. In addition, we included the percentage of female and male entrants and the total number of entrants. To facilitate interaction, in this tab filters are available for Region, State, Municipality and Courses, allowing the user to customize the visualization of the data according to their needs.



Source: The author.

7.3. Tab - "IES"

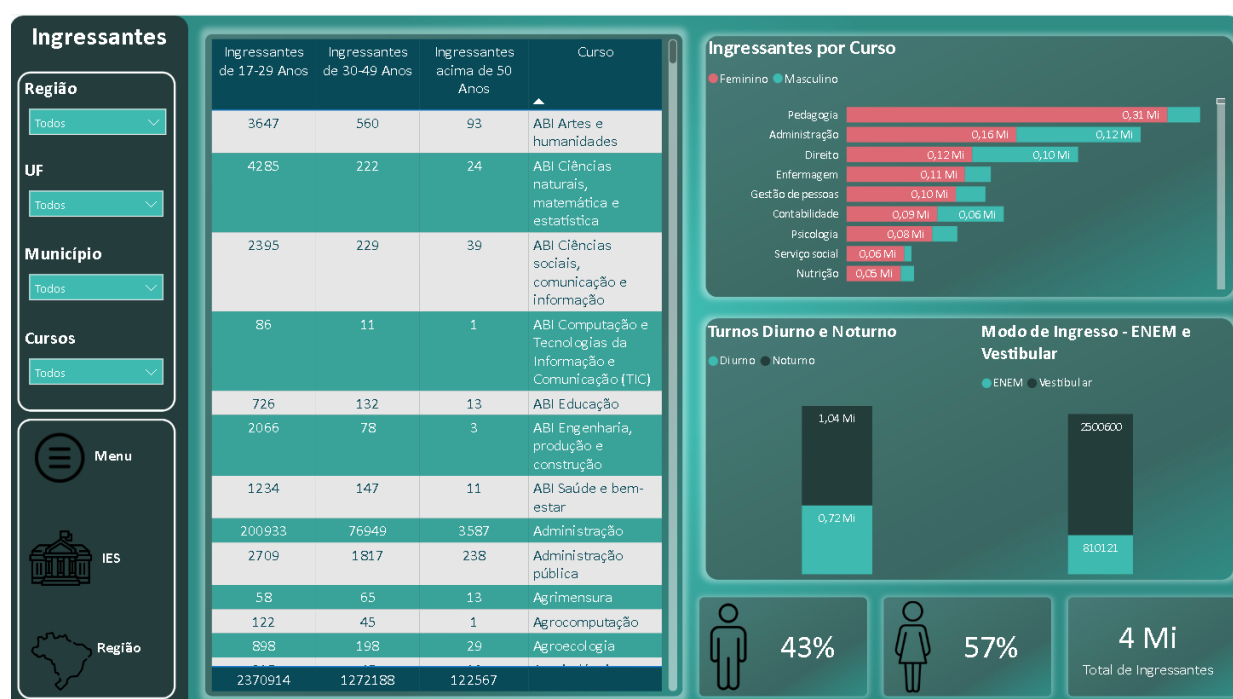
In the third tab, the analysis is presented through a map and two graphs. The map shows the number of educational institutions by region, while the first graph highlights the educational institutions with the highest number of entrants. The second graph shows the number of different courses offered by each institution. In addition, we included the percentage of female and male entrants and the total number of entrants. To improve usability, in this tab are available filters for Region, State, Municipality, Courses and HEIs, allowing the user to customize the visualization of information according to their preferences.



Source: The author.

7.4. Tab - "Joiners"

In the fourth tab, three graphs and a table are available. The first graph presents the number of freshmen per course, in addition to showing the number of female and male freshmen. The second graph shows the number of entrants coming from ENEM compared to the traditional mode of entrance exam. The third graph illustrates the number of entrants in each shift, distinguishing between day and night. The table, in turn, presents the number of entrants by age groups for each course. Also included is additional information, such as the percentage of female and male entrants and the total number of entrants. To facilitate the analysis, in this tab, the user can use filters of Region, State, Municipality and Courses.



Source: The author.

8. Measurements Calculated with DAX

8.1. Number of Distinct Courses

A measure was created using the formula "DISTINCTCOUNT" applied to the NOME_CURSO column. This approach allows you to calculate the number of distinct courses present in the dataset, providing a clear view of the diversity of degree options available.

```
Quantidade_Cursos_Distintos = DISTINCTCOUNT('Dados de Ocupação dos Cursos'[NOME_CURSO])
```

8.2. Total number of subscribers

A measurement was performed with the formula "SUM" to calculate the total number of entrants.

```
QUANT_TOTAL_INSCRITOS = SUM('Dados de Ocupação dos Cursos'[QT_ING])
```

8.3. Percentage of men enrolled

A measure was created using two variables. The first variable uses the SUM formula to calculate the total number of enrollees, while the second variable uses the SUM formula to determine the number of male entrants. With these variables, it was possible to calculate the percentage of male entrants in relation to the total number of enrollees.

```
%HOMENS_INSCRITOS =  
VAR QANT_TOTAL = SUM('Dados de Ocupação dos Cursos'[QT_ING])  
VAR QANT_TOTAL_MAS = SUM('Dados de Ocupação dos Cursos'[QT_ING_MASC])  
RETURN (QANT_TOTAL_MAS/QANT_TOTAL)
```

8.4. Percentage of women enrolled

A measure was created using two variables. The first variable uses the SUM formula to calculate the total number of enrollees, while the second variable applies the SUM formula to determine the number of female entrants. With these variables, we were able to calculate the percentage of female entrants in relation to the total number of enrollees.

```
%MULHERES_INSCRITAS =  
VAR QANT_TOTAL = SUM('Dados de Ocupação dos Cursos'[QT_ING])  
VAR QANT_TOTAL_FEM = SUM('Dados de Ocupação dos Cursos'[QT_ING_FEM])  
RETURN (QANT_TOTAL_FEM/QANT_TOTAL)
```

8.5. Number of entrants in the age group 0-29 years

A measure was created using three variables, all calculated using the SUM formula. These variables were designed to determine the total number of entrants in the age group of 0 to 29 years.

```
0-29ANOS_INSCRITOS =  
VAR FAIXA1 = SUM('Dados de Ocupação dos Cursos'[QT_ING_0_17])  
VAR FAIXA2 = SUM('Dados de Ocupação dos Cursos'[QT_ING_18_24])  
VAR FAIXA3 = SUM('Dados de Ocupação dos Cursos'[QT_ING_25_29])  
RETURN (FAIXA1+FAIXA2+FAIXA3)
```

8.6. Number of entrants in the 30-49 age group

A measure was created using three variables, all calculated using the SUM formula. These variables were designed to determine the total number of entrants in the 30-49 age group.

```
30-49ANOS_INSCRITOS =
VAR FAIXA1 = SUM('Dados de Ocupação dos Cursos'[QT_ING_30_34])
VAR FAIXA2 = SUM('Dados de Ocupação dos Cursos'[QT_ING_35_39])
VAR FAIXA3 = SUM('Dados de Ocupação dos Cursos'[QT_ING_40_49])
RETURN FAIXA1+FAIXA2+FAIXA3
```

8.7. Number of entrants over 50 years old

A measurement was performed with two variables. The variables were based on the "SUM" formula so that it is possible to find the number of entrants in the age group over 50 years.

```
>50ANOS_INSCRITOS =
VAR FAIXA1 = SUM('Dados de Ocupação dos Cursos'[QT_ING_50_59])
VAR FAIXA2 = SUM('Dados de Ocupação dos Cursos'[QT_ING_60 MAIS])
RETURN FAIXA1+FAIXA2
```

8.8. Percentage of ENEM enrollees

A measure was created using a variable, which applies the SUM formula to calculate the total number of ENEM enrollees, which is divided by the total number of enrollees.

```
%ENEM =
VAR QANT_TOTAL_ENEM = SUM('Dados de Ocupação dos Cursos'[QT_ING_ENEM])
RETURN ((QANT_TOTAL_ENEM/[QUANT_TOTAL_INSCRITOS]))
```

9. Columns Calculated with DAX

9.1. "Municipality and State"

A column calculated using the CONCATENATE function was created, which combines the information of the municipality with the acronym of the state in the table "Description of Municipalities". This approach allows you to consolidate geographic data into a single field, facilitating further analysis and visualization.

Município e Estado
Adamantina - SP
Aguaí - SP
Águas de Lindóia - SP
Águas de São Pedro - SP
Agudos - SP
Altinópolis - SP
Alumínio - SP
Alvinlândia - SP
Americana - SP

9.2. "Valid Courses"

We created a calculated column to identify "valid courses", defined as those that had a number of entrants greater than zero. For this, we use the IF function, and this column is located in the table "Course Occupancy Data". This approach allows us to filter and focus only on courses with effective student participation, optimizing our analyses.

```
Curso Válido? = IF('Dados de Ocupação dos Cursos'[QT_ING]>0,"Sim","Não")
```

10. Reach

The analysis of the data provides professionals in the area with the identification of investment opportunities and attention in specific regions, states and municipalities. This is possible due to the visualization of areas with a low number of freshmen and higher education institutions. With this information, it is feasible to direct efforts to foster higher education in these locations, contributing to the improvement of access to education and regional development.

11. End-user and level of application of the analysis

The end users of the analysis are professionals from both the public and private sectors related to the area of Education. These users can use the data strategically, allowing them to make informed decisions about investments, educational policies, and the development of initiatives aimed at improving the quality and access to higher education.

12. Technological tools implemented

The technological tools used throughout the project were:

- Microsoft Teams: For team communication and collaboration.
- Google Drive: for storing and sharing documents.
- Microsoft Word: for writing documents and reports.
- Microsoft PowerPoint: for creating presentations.
- Microsoft Excel: for data manipulation and analysis.
- Microsoft Power BI: for interactive visualization and analysis of data.
- DAX (Data Analysis Expressions): used to create calculated measures and columns, allowing for more advanced and personalized analysis of the data.

13. Future guidelines

To improve the understanding of data on higher education in Brazil, it is recommended to conduct more in-depth analyses that consider the following aspects:

1. Available Vacancies: Investigate the number of vacancies offered by higher education institutions, identifying the areas of greatest and lowest supply. This will allow for a better understanding of educational opportunities in different regions.
2. Modes of Admission: Analyze the relationship between the modes of admission (such as ENEM, vestibular, etc.) and student performance, evaluating which methods are more effective in terms of pass rate and permanence in the course.
3. Regional Distribution: Correlate information on the distribution of vacancies and entrants by state and municipality. This can help to identify regions that need greater investment in education and to propose specific interventions.

4. Demographic Segmentation: Explore data segmented by sex and race to better understand disparities in access to and retention in higher education. This analysis can highlight the need for public policies aimed at inclusion and diversity in universities.
5. Impact Evaluation: Conduct studies that evaluate the impact of educational policies on admission and permanence rates, allowing adjustments and improvements in educational management strategies.

These guidelines aim to provide valuable insights that can inform strategic decisions and public policies, contributing to the advancement of higher education in Brazil.

14. Data Source

[1] <https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/censo-da-educacao-superior>