

Universidade Cruzeiro do Sul

Nome: José Cleison de Lima

RGM: 25177591

Projeto e Aplicação de Mineração de Dados

Aplicação escolhida: Microsoft Power BI

Dataset: Compras

Download do Dataset <https://www.kaggle.com/cleisonlima/dataset-compras>

Conjunto de dados Compras:

	APC Username	APC Nome	APC Sexo	Data_Nascimento	APC Ocupação	Assinatura	Total_Gasto
1	Neaude1978	Mr bailey Glover	masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	125745
2	Maress39	Mrs isabella Fowler	feminino	04/08/1960	Arquiteto de Dados	24/03/2020	77199
3	Kehonell	Mrs Chloe Kirby	feminino	28/06/1949	Analista de RH	07/06/2020	102177
4	Thimpay	Mrs louise Clark	feminino	07/01/1995	Arquiteto de Dados	03/11/2020	23476
5	Giall1985	Mr Harrison Bowen	feminino	04/04/1987	Engenheiro de Dados	25/03/2020	62807
6	Neave1959	Mrs anna Sharp	feminino	02/02/1970	Engenheiro de Dados	18/09/2020	14004
7	Nicterims1993	Mrs Daisy Anderson	masculino	05/07/1985	Engenheiro de Dados	29/01/2021	241781
8	Fouldlairity	Ms zoe Peacock	masculino	03/06/1964	Analista Financeiro	04/12/2020	5195
9	Ancomettiody	Mrs Libby Williams	feminino	04/09/1974	Estatístico	21/11/2020	13073
10	Onswis	Ms Katie Mmtcalfe	feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	177872
11	Majeappre	Mr Luca Vaughan	feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	103474
12	Theired	Mr Henry Haynes	masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	132004
13	Botter	Mr Noah Walton	feminino	09/03/1991	Web Designer	26/07/2022	346472
14				null		null	null

Limpeza e Transformação dos dados:

1. Remoção de Linha em branco

- Página inicial – Reduzir Linhas – Remover Linhas – Remover Linhas Inferiores – Número 1 – Ok

Antes:

	APC Username	APC Nome	APC Sexo	Data_Nascimento	APC Ocupação	Assinatura	Total_Gasto
1	Neaude1978	Mr bailey Glover	masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	125745
2	Maress39	Mrs isabella Fowler	feminino	04/08/1960	Arquiteto de Dados	24/03/2020	77199
3	Kehonell	Mrs Chloe Kirby	feminino	28/06/1949	Analista de RH	07/06/2020	102177
4	Thimpay	Mrs louise Clark	feminino	07/01/1995	Arquiteto de Dados	03/11/2020	23476
5	Giall1985	Mr Harrison Bowen	feminino	04/04/1987	Engenheiro de Dados	25/03/2020	62807
6	Neave1959	Mrs anna Sharp	feminino	02/02/1970	Engenheiro de Dados	18/09/2020	14004
7	Nicterims1993	Mrs Daisy Anderson	masculino	05/07/1985	Engenheiro de Dados	29/01/2021	241781
8	Fouldlairity	Ms zoe Peacock	masculino	03/06/1964	Analista Financeiro	04/12/2020	5195
9	Ancomettiody	Mrs Libby Williams	feminino	04/09/1974	Estatístico	21/11/2020	13073
10	Onswis	Ms Katie Mmtcalfe	feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	177872
11	Majeappre	Mr Luca Vaughan	feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	103474
12	Theired	Mr Henry Haynes	masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	132004
13	Botter	Mr Noah Walton	feminino	09/03/1991	Web Designer	26/07/2022	346472
14				null		null	null

Depois:

	📅👤 Username	👤 Nome	👤 Sexo	📅 Data_Nascimento	👤 Ocupação	📅 Assinatura	123 Total_Gasto
1	Neaude1978	Mr bailey Glover	masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	125745
2	Mares39	Mrs isabella Fowler	feminino	04/08/1960	Arquiteto de Dados	24/03/2020	77199
3	Kehonell	Mrs Chloe Kirby	feminino	28/06/1949	Analista de RH	07/06/2020	102177
4	Thimpay	Mrs louise Clark	feminino	07/01/1995	Arquiteto de Dados	03/11/2020	23476
5	Giall1985	Mr Harrison Bowen	feminino	04/04/1987	Engenheiro de Dados	25/03/2020	62807
6	Neave1959	Mrs anna Sharp	feminino	02/02/1970	Engenheiro de Dados	18/09/2020	14004
7	Nicterims1993	Mrs Daisy Anderson	masculino	05/07/1985	Engenheiro de Dados	29/01/2021	241781
8	Fouldlarity	Ms zoe Peacock	masculino	03/06/1964	Analista Financeiro	04/12/2020	5195
9	Ancometdiody	Mrs Libby Williams	feminino	04/09/1974	Estatístico	21/11/2020	13073
10	Onswis	Ms Katie Mmtcalfe	feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	177872
11	Majeappre	Mr Luca Vaughan	feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	103474
12	Theired	Mr Henry Haynes	masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	132004
13	Botter	Mr Noah Walton	feminino	09/03/1991	Web Designer	26/07/2022	346472

2. Remoção do Atributo Username:

- Lado direito do mouse em Username - Remove

Antes:

	📅👤 Nome	👤 Sexo	📅 Data_Nascimento	👤 Ocupação	📅 Assinatura	123 Total_Gasto
1	Mr bailey Glover	masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	125745
2	Mrs isabella Fowler	feminino	04/08/1960	Arquiteto de Dados	24/03/2020	77199
3	Mrs Chloe Kirby	feminino	28/06/1949	Analista de RH	07/06/2020	102177
4	Mrs louise Clark	feminino	07/01/1995	Arquiteto de Dados	03/11/2020	23476
5	Mr Harrison Bowen	feminino	04/04/1987	Engenheiro de Dados	25/03/2020	62807
6	Mrs anna Sharp	feminino	02/02/1970	Engenheiro de Dados	18/09/2020	14004
7	Mrs Daisy Anderson	masculino	05/07/1985	Engenheiro de Dados	29/01/2021	241781
8	Ms zoe Peacock	masculino	03/06/1964	Analista Financeiro	04/12/2020	5195
9	Mrs Libby Williams	feminino	04/09/1974	Estatístico	21/11/2020	13073
10	Ms Katie Mmtcalfe	feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	177872
11	Mr Luca Vaughan	feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	103474
12	Mr Henry Haynes	masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	132004
13	Mr Noah Walton	feminino	09/03/1991	Web Designer	26/07/2022	346472

Depois:

	📅👤 Nome	👤 Sexo	📅 Data_Nascimento	👤 Ocupação	📅 Assinatura	123 Total_Gasto
1	Mr bailey Glover	masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	125745
2	Mrs isabella Fowler	feminino	04/08/1960	Arquiteto de Dados	24/03/2020	77199
3	Mrs Chloe Kirby	feminino	28/06/1949	Analista de RH	07/06/2020	102177
4	Mrs louise Clark	feminino	07/01/1995	Arquiteto de Dados	03/11/2020	23476
5	Mr Harrison Bowen	feminino	04/04/1987	Engenheiro de Dados	25/03/2020	62807
6	Mrs anna Sharp	feminino	02/02/1970	Engenheiro de Dados	18/09/2020	14004
7	Mrs Daisy Anderson	masculino	05/07/1985	Engenheiro de Dados	29/01/2021	241781
8	Ms zoe Peacock	masculino	03/06/1964	Analista Financeiro	04/12/2020	5195
9	Mrs Libby Williams	feminino	04/09/1974	Estatístico	21/11/2020	13073
10	Ms Katie Mmtcalfe	feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	177872
11	Mr Luca Vaughan	feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	103474
12	Mr Henry Haynes	masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	132004
13	Mr Noah Walton	feminino	09/03/1991	Web Designer	26/07/2022	346472

3. Formato de data incorreto:

- Clic na opção ao lado (ABC) – Usando a localidade – Tipo de Dados: Data – Localidade (Português)

Antes:

	ABC Nome	ABC Sexo	ABC Data_Nascimento	ABC Ocupação	ABC Assinatura	Total_Gasto
1	Mr bailey Glover	masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	125745
2	Mrs isabella Fowler	feminino	04/08/1960	Arquiteto de Dados	24/03/2020	77199
3	Mrs Chloe Kirby	feminino	28/06/1949	Analista de RH	07/06/2020	102177
4	Mrs louise Clark	feminino	07/01/1995	Arquiteto de Dados	03/11/2020	23476
5	Mr Harrison Bowen	feminino	04/04/1987	Engenheiro de Dados	25/03/2020	62807
6	Mrs anna Sharp	feminino	02/02/1970	Engenheiro de Dados	18/09/2020	14004
7	Mrs Daisy Anderson	masculino	05/07/1985	Engenheiro de Dados	29/01/2021	241781
8	Ms zoe Peacock	masculino	03/06/1964	Analista Financeiro	04/12/2020	5195
9	Mrs Libby Williams	feminino	04/09/1974	Estatístico	21/11/2020	13073
10	Ms Katie Mmtcalfe	feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	177872
11	Mr Luca Vaughan	feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	103474
12	Mr Henry Haynes	masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	132004
13	Mr Noah Walton	feminino	09/03/1991	Web Designer	26/07/2022	346472

Depois:

	ABC Nome	ABC Sexo	Data_Nascimento	ABC Ocupação	Assinatura	Total_Gasto
1	Mr bailey Glover	masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	125745
2	Mrs isabella Fowler	feminino	04/08/1960	Arquiteto de Dados	24/03/2020	77199
3	Mrs Chloe Kirby	feminino	28/06/1949	Analista de RH	07/06/2020	102177
4	Mrs louise Clark	feminino	07/01/1995	Arquiteto de Dados	03/11/2020	23476
5	Mr Harrison Bowen	feminino	04/04/1987	Engenheiro de Dados	25/03/2020	62807
6	Mrs anna Sharp	feminino	02/02/1970	Engenheiro de Dados	18/09/2020	14004
7	Mrs Daisy Anderson	masculino	05/07/1985	Engenheiro de Dados	29/01/2021	241781
8	Ms zoe Peacock	masculino	03/06/1964	Analista Financeiro	04/12/2020	5195
9	Mrs Libby Williams	feminino	04/09/1974	Estatístico	21/11/2020	13073
10	Ms Katie Mmtcalfe	feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	177872
11	Mr Luca Vaughan	feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	103474
12	Mr Henry Haynes	masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	132004
13	Mr Noah Walton	feminino	09/03/1991	Web Designer	26/07/2022	346472

4. Atributo **Nome** com os Dados com excesso de espaços e Iniciais dos nomes sem as devidas letras maiúsculas:

- Para a remoção fui na parte de transformação – Formato – Cortar (Trim) – Limpar.
- Depois substituir valores – Lado direito do mouse na variável – Substituir valores.

Antes

Nome
Mr bailey Glover
Mrs isabella Fowler
Mrs Chloe Kirby
Mrs louise Clark
Mr Harrison Bowen
Mrs anna Sharp
Mrs Daisy Anderson
Ms zoe Peacock
Mrs Libby Williams
Ms Katie Mmtcalfe
Mr Luca Vaughan
Mr Henry Haynes
Mr Noah Walton

Depois

Nome
Mr bailey Glover
Mrs isabella Fowler
Mrs Chloe Kirby
Mrs louise Clark
Mr HarrisonBowen
Mrs annaSharp
Mrs Daisy Anderson
Ms zoe Peacock
Mrs Libby Williams
Ms Katie Mmtcalfe
Mr Luca Vaughan
Mr Henry Haynes
Mr Noah Walton

5. Atributo **Total Gasto** no formato de número inteiro:

- Painel de controle do Windows – Barra de Busca – Região – Formato: Português
- Ícone de formato – Usando a localidade – Tipo de Dados: Decimal Fixo – Localidade: Português Brasil

Antes

1.2 Total_Gasto
1257.45
771.99
1021.77
234.76
628.07
140.04
2417.81
51.95
130.73
1778.72
1034.74
1320.04
3464.72

Depois

\$ Total_Gasto
1.257,45
771,99
1.021,77
234,76
628,07
140,04
2.417,81
51,95
130,73
1.778,72
1.034,74
1.320,04
3.464,72

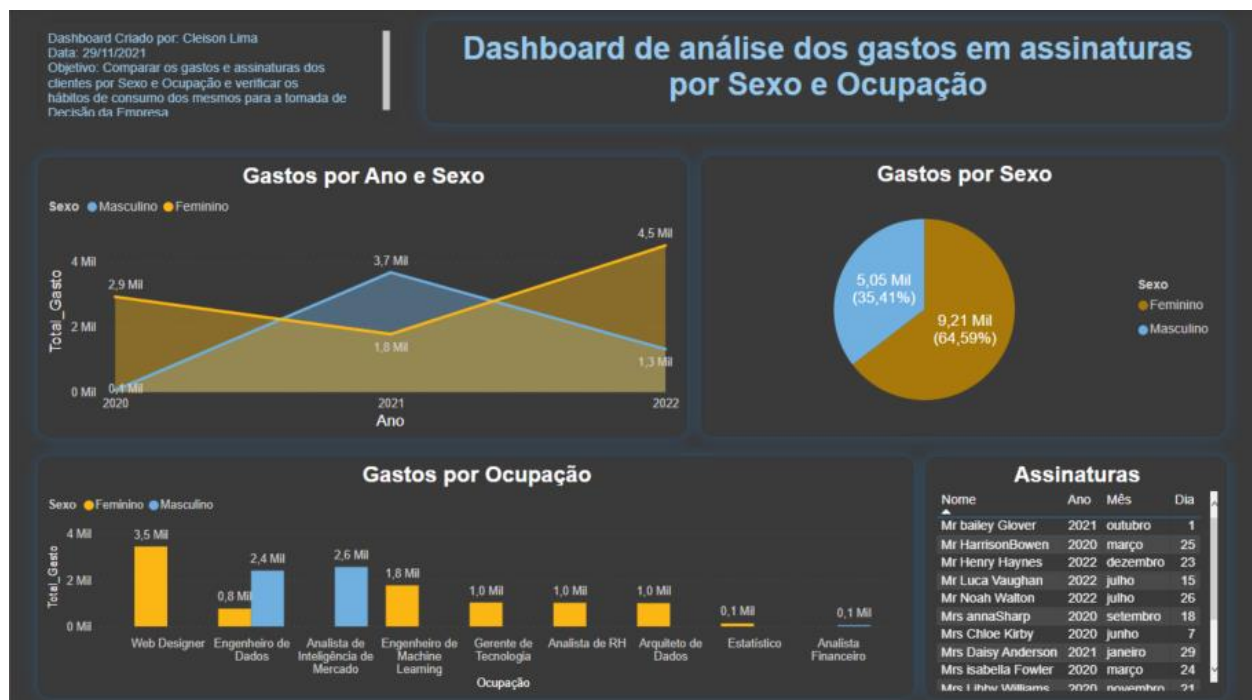
Fim da limpeza e Transformação dos Dado.

Resultado:

	A _C Nome	A _C Sexo	Data_Nascimento	A _C Ocupação	Assinatura	\$ Total_Gasto
1	Mr bailey Glover	Masculino	01/12/1979	Analista de Inteligência de Mercado	01/10/2021	1.257,45
2	Mrs isabella Fowler	Feminino	04/08/1960	Arquiteto de Dados	24/03/2020	771,99
3	Mrs Chloe Kirby	Feminino	28/06/1949	Analista de RH	07/06/2020	1.021,77
4	Mrs louise Clark	Feminino	07/01/1995	Arquiteto de Dados	03/11/2020	234,76
5	Mr HarrisonBowen	Feminino	04/04/1987	Engenheiro de Dados	25/03/2020	628,07
6	Mrs annaSharp	Feminino	02/02/1970	Engenheiro de Dados	18/09/2020	140,04
7	Mrs Daisy Anderson	Masculino	05/07/1985	Engenheiro de Dados	29/01/2021	2.417,81
8	Ms zoe Peacock	Masculino	03/06/1964	Analista Financeiro	04/12/2020	51,95
9	Mrs Libby Williams	Feminino	04/09/1974	Estatístico	21/11/2020	130,73
10	Ms Katie Mmtcalfe	Feminino	01/07/1983	Engenheiro de Machine Learning	27/02/2021	1.778,72
11	Mr Luca Vaughan	Feminino	03/01/1971	Gerente de Tecnologia	15/07/2022	1.034,74
12	Mr Henry Haynes	Masculino	06/09/1951	Analista de Inteligência de Mercado	23/12/2022	1.320,04
13	Mr Noah Walton	Feminino	09/03/1991	Web Designer	26/07/2022	3.464,72

Principal objetivo da visualização dos dados: Encontrar padrões de consumo dos clientes classificando os dados por Gastos, Sexo e Ocupação e plotando em um Dashboard interativo para a tomada de decisão.

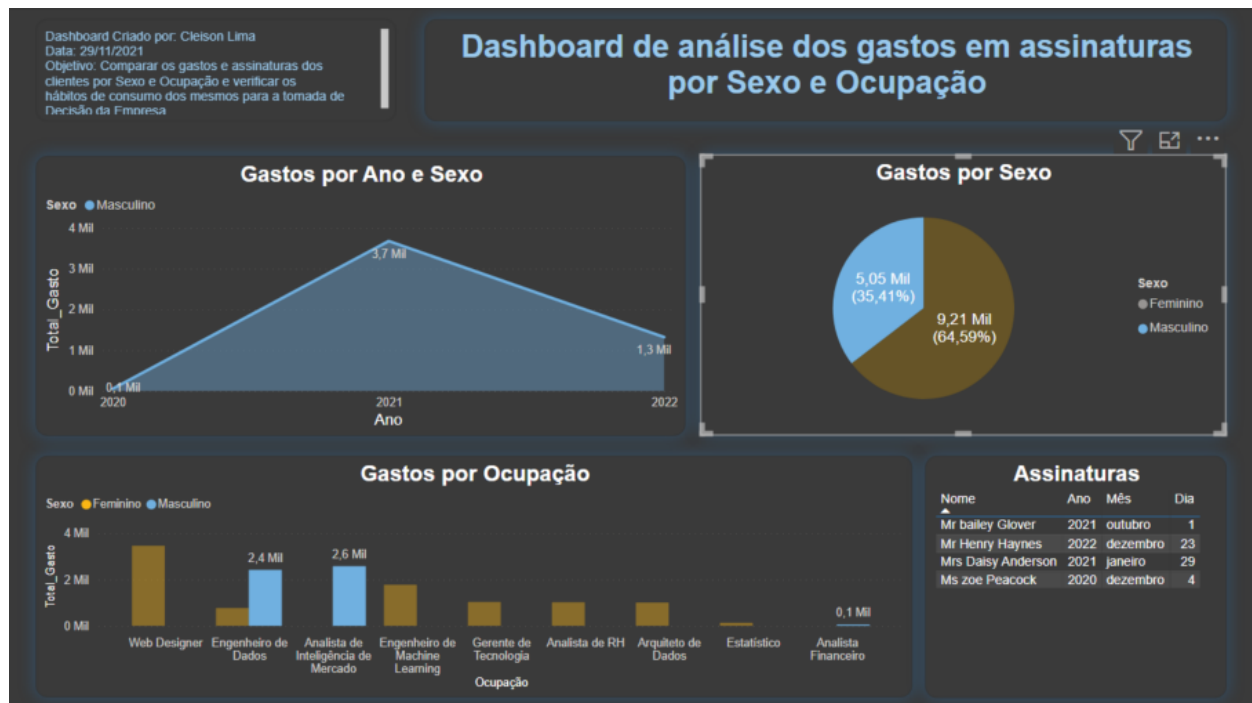
Visualização dos Dados:



Clicando nos dados do Sexo Masculino em **Azul** e verificando a interatividade do Dashboard:

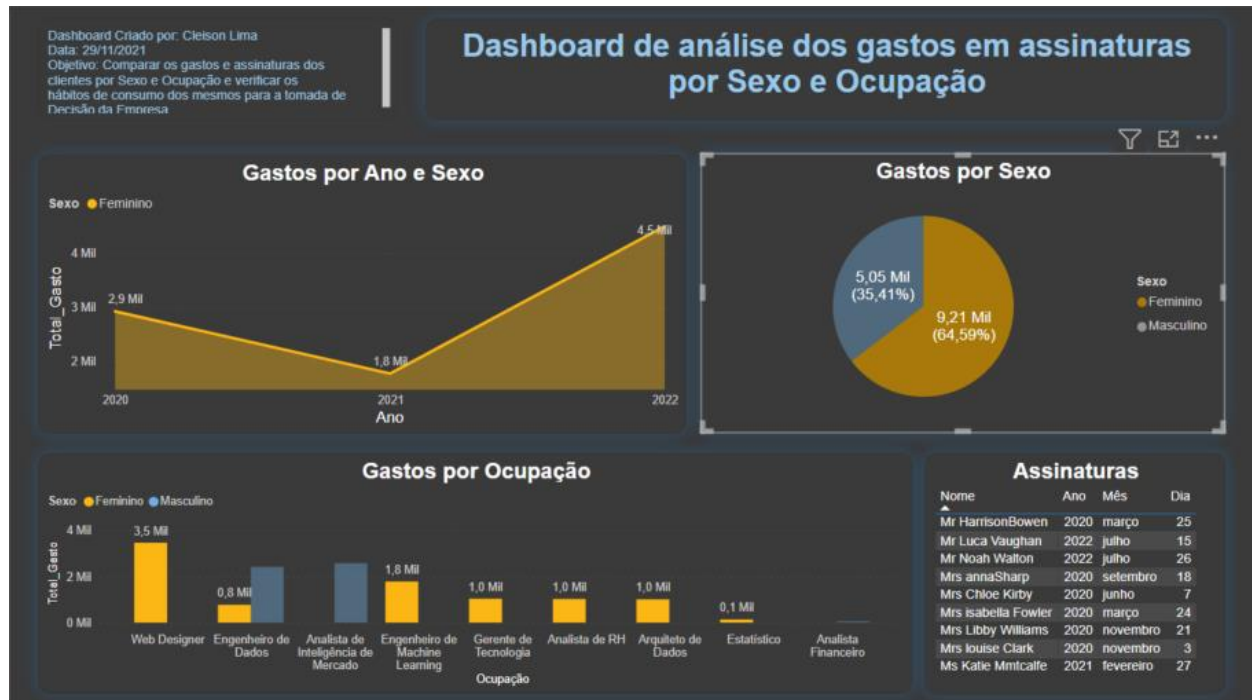
Alto consumo no ano 2021 e queda em 2022.

Os que mais consomem por Ocupação: Analistas de Inteligência de Mercado e Engenheiros de dados



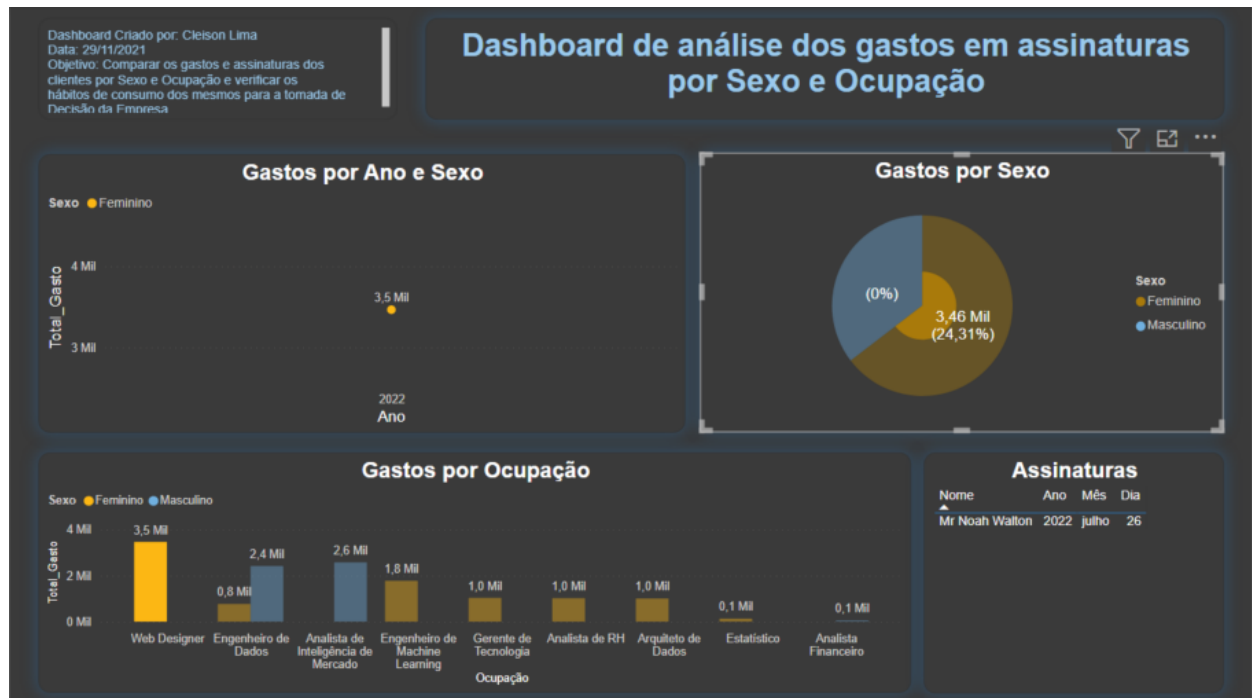
Clicando nos dados do Sexo Feminino em **Amarelo** e verificando a interatividade do Dashboard:

Baixo consumo no ano 2021 e alto em 2022.



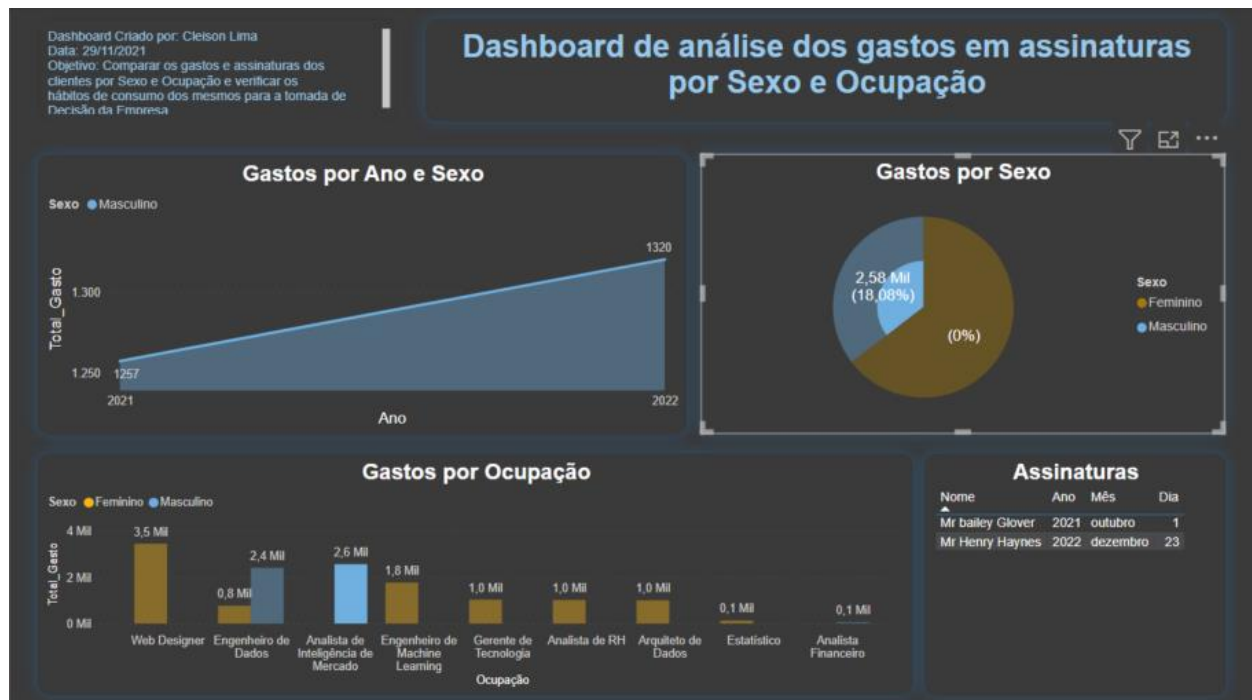
Clicando nos dados de Gasto por ocupação do sexo Feminino em **Amarelo** e verificando a interatividade do Dashboard:

Pessoas com a profissão de Web Design do sexo Feminino gastam mais que os demais.



Clicando nos dados de Gasto por ocupação do sexo Masculino em **Azul** e verificando a interatividade do Dashboard:

Pessoas com a profissão de Analista de Inteligência de Mercado do sexo Masculino gastam bem mais que os demais ficando abaixo das mulheres da profissão da profissão de Web Designer.



Considerações finais:

Com isso conseguimos ver quais são os públicos alvos por sexo e as que mais são propensas a realizarem mais assinaturas.

De acordo com o plano de negócio da empresa muitos outros fatores serão de máxima importância para colocarmos como ponto focal de vendas para o público do sexo masculino e feminino.