**Relatório atividade EBAC**

Para a primeira etapa selecionei a fonte de dados do Kaggle, referente ao comportamento dos consumidores em relação à compras feitas pelo E-comerce, disponível no link abaixo:

<https://www.kaggle.com/datasets/salahuddinahmedshuvo/ecommerce-consumer-behavior-analysis-data>

Através dos dados fornecidos, busquei entender um pouco melhor como está o comportamento de compra avaliando diversas diferenças, como idade, escolaridade, frequência de uso. Para entender um pouco melhor se esse tipo de comercio está mais amigável para toda a população.

Os dados utilizados foram dados estruturados, apenas em uma tabela. Inicialmente com 28 colunas, mas novas colunas foram adicionadas para análises especificas. Como dito inicialmente, dado foi retirado do kaggle, e fizemos todo tratamento através do PyCharm, utilizando python.

Segue o código abaixo:

import pandas as pd  
from sklearn.preprocessing import LabelEncoder  
  
  
pd.set\_option('display.width', None)  
pd.set\_option('display.max\_colwidth', None)  
  
df = pd.read\_csv(r'C:\Users\Felip\Downloads\Ecommerce\_Consumer\_Behavior\_Analysis\_Data.csv')  
print(df.head())  
  
print(df.shape)  
print(df.dtypes)  
print('Valores Nulos:\n', df.isnull().sum())  
  
df\_tratado = df.fillna('Unknown')  
print('Valores Nulos:\n', df\_tratado.isnull().sum())  
  
ensino = df\_tratado['Education\_Level'].unique()  
print(ensino)  
  
education\_ordem = {"Bachelor's": 1, "High School": 2, "Master's": 3}  
df\_tratado['Education\_Level\_Ord'] = df['Education\_Level'].map(education\_ordem)  
  
label\_encoder = LabelEncoder()  
df\_tratado['Gender\_cod'] = label\_encoder.fit\_transform(df['Gender'])  
df\_tratado['Marital\_Status\_cod'] = label\_encoder.fit\_transform(df['Marital\_Status'])  
  
df\_tratado['Purchase\_Amount'] = df\_tratado['Purchase\_Amount'].str.replace('$', '', regex=False)  
df\_tratado['Purchase\_Amount'] = df\_tratado['Purchase\_Amount'].astype(float)  
  
  
bins = [17, 23, 29, 35, 41, 47, 60]  
labels = ['18-23', '24-29', '30-35', '36-41', '42-47', '48-60']  
df\_tratado['Age\_range'] = pd.cut(df['Age'], bins=bins, labels=labels)  
  
print(df\_tratado.dtypes)  
print(df\_tratado['Purchase\_Amount'].sum())  
print(df\_tratado.head())  
  
  
print('Correlação: \n',df\_tratado[['Purchase\_Amount', 'Age','Gender\_cod', 'Marital\_Status\_cod','Education\_Level\_Ord', 'Frequency\_of\_Purchase', 'Time\_Spent\_on\_Product\_Research(hours)', 'Customer\_Satisfaction', 'Return\_Rate']].corr())  
  
df\_tratado.to\_csv('Df\_tratado.csv', index=False)

O arquivo será encaminhado junto a atividade.

No último passo foi feita a apresentação no Looker Studio, como solicitado, abaixo segue o link para acesso do relatório.

<https://lookerstudio.google.com/s/uVfQlcmiSCo>

Conclusões:

Através dos insights, conseguimos perceber que E-comerce está cada mais aceitável para a população de forma geral. Antigamente era comum esse método ser utilizados apenas por pessoas mais novas, com mais intimidade com o mundo digital. Mas podemos ver que através desses dados que essa diferença está cada vez menor, em muitos casos até irrelevantes. Não importa se é homem ou mulher, a idade ou escolaridade. Hoje em dia todos tem valores bem próximos de gastos com e-comerce.