import numpy as np

import pandas as pd

pd.set\_option("display.precision", 2)

df = pd.read\_csv("../input/Clothing.csv")

df.head()

print(df.shape)

print(df.columns)

print(df.info())

df["Churn"] = df["Churn"].astype("int64")

df.describe()

df.describe(include=["object", "bool"])

df["Churn"].value\_counts()

df["Churn"].value\_counts(normalize=True)

df.sort\_values(by="Total Cloth Price", ascending=False).head()

df.sort\_values(by=["Churn", "Total Cloth Price"], ascending=[True, False]).head()

df["Churn"].mean()

df[df["Churn"] == 1].mean()

df[df["Churn"] == 1]["Total Cloth Price"].mean()

df[(df["Churn"] == 0) & (df["Cloth Price"] == "No")]["Total Cloth Price"].max()

df.loc[0:5, "State":"Area code"]

df.iloc[0:5, 0:3]

df[-1:]

df.apply(np.max)

df[df["State"].apply(lambda state: state[0] == "W")].head()

d = {"No": False, "Yes": True}

df["Cloth price "] = df["Cloth price "].map(d)

df.head()

df.groupby(by=grouping\_columns)[columns\_to\_show].function()

columns\_to\_show = ["Cloth price ", " Cloth price ", " Cloth price "]

df.groupby(["Churn"])[columns\_to\_show].describe(percentiles=[])

columns\_to\_show = ["Cloth price ", " Cloth price ", " Cloth price "]

df.groupby(["Churn"])[columns\_to\_show].agg([np.mean, np.std, np.min, np.max])

pd.crosstab(df["Churn"], df["Cloth price"])

pd.crosstab(df["Churn"], df["Cloth price "], normalize=True)

df.pivot\_table(

["Cloth price ", " Cloth price ", "Cloth price "],

["Area code"],

aggfunc="mean",

)

total\_calls = (

df["Cloth price "]

+ df["Total eve calls"]

+ df["Total night calls"]

+ df["Total intl calls"]

)

df.insert(loc=len(df.columns), column="Total calls", value=total\_calls)

*# loc parameter is the number of columns after which to insert the Series object*

*# we set it to len(df.columns) to paste it at the very end of the dataframe*

df.head()

df["Total charge"] = (

df["Cloth price "]

+ df["Total eve charge"]

+ df["Total night charge"]

+ df["Total area charge"]

)

df.head()

*# get rid of just created columns*

df.drop(["Cloth price ", " Cloth price "], axis=1, inplace=True)

*# and here’s how you can delete rows*

df.drop([1, 2]).head()

pd.crosstab(df["Churn"], df["Cloth price "], margins=True)

*# some imports to set up plotting*

import matplotlib.pyplot as plt

*# pip install seaborn*

import seaborn as sns

*# Graphics in retina format are more sharp and legible*

%config InlineBackend.figure\_format = 'retina'

sns.countplot(x=" Cloth price ", hue="Churn", data=df);

pd.crosstab(df["Churn"], df["Cloth price margins=True)

sns.countplot(x=" Cloth price ", hue="Churn", data=df);

df["Cloth price "] = (df["Cloth price "] > 3).astype("int")

pd.crosstab(df["Cloth price "], df["Churn"], margins=True)

sns.countplot(x=" Cloth price ", hue="Churn", data=df);

pd.crosstab(df["Cloth price "] & df["Cloth price "], df["Churn"])