Heroes of Pymoli

# Dependencies and Setup

import pandas as pd

import numpy as np

# File to Load (Remember to Change These)

file\_to\_load = "purchase\_data.csv"

# Read Purchasing File and store into Pandas data frame

purchases\_df = pd.read\_csv(file\_to\_load)

purchases\_df.head()

p\_count = len(purchases\_df["SN"].unique())

p\_count\_df = pd.DataFrame({"Player Count":[p\_count]})

p\_count\_df

#Total Number of Purchases

#Total Revenue

unique\_items = len(purchase\_data["Item Name"].unique())

avg\_price = purchases\_df["Price"].mean()

purch\_tot = purchases\_df["Price"].sum()

purch\_num = len(purchases\_df["Purchase ID"].unique())

summ\_df = pd.DataFrame({"Item Name":[unique\_items],

"Average Price":[avg\_price],

"Total Revenue":[purch\_tot],

"Total Number of Purchases":[purch\_num]})

summ\_df.round({"Average Price":2})

#get rid of duplicate Usernames

unique\_df = purchases\_df.drop\_duplicates(subset="SN")

#new data frame with counts of genders

gender\_df = unique\_df["Gender"].value\_counts()

#made variables and iloc to get particular gender counts

male\_num = gender\_df.iloc[0]

female\_num = gender\_df.iloc[1]

undisc\_num = gender\_df.iloc[2]

#get percentages by dividing by total(p\_count)

male\_pct = round((gender\_df.iloc[0]/p\_count)\*100,1)

female\_pct = round((gender\_df.iloc[1]/p\_count)\*100,1)

undisc\_pct = round((gender\_df.iloc[2]/p\_count)\*100,1)

summ\_gender = pd.DataFrame({"Gender":["Male", "Female","Undisclosed"],

"Counts":[male\_num,female\_num,undisc\_num],

"Percent %":['{:,.2f}%'.format(male\_pct),'{:,.2f}%'.format(female\_pct),'{:,.2f}%'.format(undisc\_pc

gen\_counts = purchases\_df["Gender"].value\_counts()

#purchase counts

male\_pcount = gen\_counts.iloc[0]

female\_pcount = gen\_counts.iloc[1]

undisc\_pcount = gen\_counts.iloc[2]

# Total spent by gender

total\_purch = purchases\_df.groupby(['Gender']).sum()

male\_tot\_price = round(total\_purch.iloc[1,3],2)

female\_tot\_price = round(total\_purch.iloc[0,3],2)

undisc\_tot\_price = round(total\_purch.iloc[2,3],2)

# Average by gender

male\_avg = round((male\_tot\_price/male\_pcount),2)

female\_avg = round((female\_tot\_price/female\_pcount),2)

undisc\_avg = round((undisc\_tot\_price/undisc\_pcount),2)

#Average by individual by gender

male\_ind\_avg = round((male\_tot\_price/male\_num),2)

fem\_ind\_avg = round((female\_tot\_price/female\_num),2)

undisc\_ind\_avg = round((undisc\_tot\_price/undisc\_num),2)

summ\_gender\_price = pd.DataFrame({"Gender":["Male", "Female","Undisclosed"],

"Purchase Counts":[male\_pcount,female\_pcount,undisc\_pcount],

"Average Purchase Price":['${:,.2f}'.format(male\_avg),'${:,.2f}'.format(female\_avg),'${:,.2f}'.format(undisc\_avg)],

"Total Purchase Value":['${:,.2f}'.format(male\_tot\_price),'${:,.2f}'.format(female\_tot\_price),'${:,.2f}'.format(undisc\_tot\_price)],

"Avg Purchase Per Person":['${:,.2f}'.format(male\_ind\_avg),'${:,.2f}'.format(fem\_ind\_avg),'${:,.2f}'.format(undisc\_ind\_avg)]})

summ\_gender\_price