main

September 7, 2021

[1]:
```
import os
import pandas as pd
filepath = os.path.join ('Resources', 'purchase_data.csv')
purchase_df = pd.read_csv(filepath)
purchase_df.head()
```

[1]:

	Purchase ID	SN	Age	Gender	Item ID	Item Name	Price
0	0	Lisim78	20	Male	108	Extraction, Quickblade Of Trembling Hands	3.53
1	1	Lisovynya38	40	Male	143	Frenzied Scimitar	1.56
2	2	Ithergue48	24	Male	92	Final Critic	4.88
3	3	Chamassasya86	24	Male	100	Blindscythe	3.27
4	4	Iskosia90	23	Male	131	Fury	1.44

[2]:
```
total_players = len(purchase_df['SN'].unique())

total_players_result = pd.DataFrame([{'Total Players': total_players}])
total_players_result
```

[2]:

	Total Players
0	576

[3]:
```
numbers_of_unique_items = len(purchase_df['Item ID'].unique())
average_price = purchase_df['Price'].mean()
average_price = round(average_price,2)                    #Keep 2 decimals
average_price = "${:.2f}".format(average_price)           #Add the Dollar sign
number_of_purchases = len(purchase_df.index)
total_revenue = purchase_df['Price'].sum()
total_revenue = "${:.2f}".format(total_revenue)              #Add the Dollar␣
↪sign
```

```
purchasing_analysis_total = pd.DataFrame({'Number of Unique Items':␣
↪[numbers_of_unique_items],
                                          'Average Price': [average_price],
                                          'Number of Purchases':␣
↪[number_of_purchases],
                                          'Total Revenue': [total_revenue]})
purchasing_analysis_total
```

[3]:

	Number of Unique Items	Average Price	Number of Purchases	Total Revenue
0	179	$3.05	780	$2379.77

[4]:
```
unique_purchase_df = purchase_df.drop_duplicates(subset=['SN'])        #Drop␣
↪the rows with the same SN to get the unique players
number_of_male = len(unique_purchase_df.loc[purchase_df['Gender'] == "Male",:])
number_of_female = len(unique_purchase_df.loc[unique_purchase_df['Gender'] ==␣
↪'Female', :])
number_of_other = len(unique_purchase_df.loc[unique_purchase_df['Gender'] ==␣
↪'Other / Non-Disclosed', :])
percentage_of_male = (number_of_male/total_players)
percentage_of_male = "{:.2%}".format(percentage_of_male)     #Adding the dollar␣
↪sign
percentage_of_female = (number_of_female/total_players)
percentage_of_female = "{:.2%}".format(percentage_of_female)
percentage_of_other = (number_of_other/total_players)
percentage_of_other = "{:.2%}".format(percentage_of_other)

gender_demographics = pd.DataFrame({'Gender':['Male','Female','Other /␣
↪Non-Disclosed'],
                                    'Total Count':[number_of_male,␣
↪number_of_female,number_of_other],
                                    'Percentage of Players':␣
↪[percentage_of_male,percentage_of_female,percentage_of_other]})
gender_demographics.set_index('Gender')
```

[4]:

Gender	Total Count	Percentage of Players
Male	484	84.03%
Female	81	14.06%
Other / Non-Disclosed	11	1.91%

[5]:
```
grouped_purchase_df = purchase_df.groupby(['Gender'])

purchase_count = grouped_purchase_df['Purchase ID'].count()                ␣
↪#Purchase count per gender
average_price = list(grouped_purchase_df['Price'].mean())                  ␣
↪#Avg purchase price per gender
```

```
total_purchase_value = purchase_count * average_price
#Total purchase value
total_purchase_value_list = list(purchase_count * average_price)
avg_total_purchase_per_person = list(total_purchase_value/
(number_of_female,number_of_male,number_of_other))
grouped_purchase_df = pd.DataFrame(grouped_purchase_df.size().reset_index(name
= "Purchase Count"))
grouped_purchase_df ['Average Purchase Price'] = average_price
grouped_purchase_df ['Average Purchase Price'] = grouped_purchase_df ['Average
Purchase Price'].map("${:.2f}".format)
grouped_purchase_df ['Total Purchase Value'] = total_purchase_value_list
grouped_purchase_df ['Total Purchase Value'] = grouped_purchase_df ['Total
Purchase Value'].map("${:.2f}".format)
grouped_purchase_df ['Avg Total Purchase per Person'] =
avg_total_purchase_per_person
grouped_purchase_df ['Avg Total Purchase per Person'] = grouped_purchase_df
['Avg Total Purchase per Person'].map("${:.2f}".format)
grouped_purchase_df.set_index('Gender')
```

[5]:

Gender	Purchase Count	Average Purchase Price	Total Purchase Value	Avg Total Purchase per Person
Female	113	$3.20	$361.94	$4.47
Male	652	$3.02	$1967.64	$4.07
Other / Non-Disclosed	15	$3.35	$50.19	$4.56

[6]:
```
unique_sn_df = purchase_df.drop_duplicates(subset = 'SN', keep = 'last').copy()
bins = [0,9.9,14,19,24,29,34,39,100]
group_names = ["<10", "10-14", '15-19','20-24','25-29','30-34','35-39','40+']
age = list(unique_sn_df ['Age'])
pd.cut(unique_sn_df['Age'],bins,labels=group_names)
unique_sn_df['Age Range'] = pd.cut(unique_sn_df['Age'],bins,labels=group_names)
age_df = unique_sn_df.groupby('Age Range')
total_count = list(age_df['Age Range'].count())
age_range_percentage = [age/len(unique_sn_df.index)*100 for age in total_count]
age_range_df = pd.DataFrame ({'':group_names,'Total Count':total_count,
'Percentage of Players':age_range_percentage})
age_range_df['Percentage of Players'] = age_range_df['Percentage of Players'].
map("{:.2f}%".format)
age_range_df.reset_index()
age_range_df.set_index('')
```

[6]:

	Total Count	Percentage of Players
<10	17	2.95%
10-14	22	3.82%
15-19	107	18.58%
20-24	258	44.79%
25-29	77	13.37%
30-34	52	9.03%
35-39	31	5.38%
40+	12	2.08%

[7]:
```
purchase_df['Age Ranges'] = pd.cut(purchase_df['Age'],bins,labels=group_names).
 ↪copy()
grouped_purchase_df = purchase_df.groupby('Age Ranges')
purchase_count = grouped_purchase_df['SN'].count()
average_purchase_price = grouped_purchase_df['Price'].mean()
totol_purchase_value = purchase_count * average_purchase_price
average_total_purchase_per_person = grouped_purchase_df['Price'].sum()/
 ↪age_df['Age Range'].count()
purchasing_analysis_age_df = pd.DataFrame ({'Purchase Count':purchase_count,
                                            'Average Purchase Price':
 ↪average_purchase_price,
                                            'Total Purchase Value':
 ↪totol_purchase_value,
                                            'Avg Total Purchase per Person':
 ↪average_total_purchase_per_person})
purchasing_analysis_age_df['Average Purchase Price'] =␣
 ↪purchasing_analysis_age_df['Average Purchase Price'].map("${:.2f}".format)
purchasing_analysis_age_df['Total Purchase Value']=␣
 ↪purchasing_analysis_age_df['Total Purchase Value'].map("${:.2f}".format)
purchasing_analysis_age_df['Avg Total Purchase per Person']=␣
 ↪purchasing_analysis_age_df['Avg Total Purchase per Person'].map("${:.2f}".
 ↪format)
purchasing_analysis_age_df
```

[7]:

Age Ranges	Purchase Count	Average Purchase Price	Total Purchase Value	Avg Total Purchase per Person
<10	23	$3.35	$77.13	
10-14	28	$2.96	$82.78	
15-19	136	$3.04	$412.89	
20-24	365	$3.05	$1114.06	
25-29	101	$2.90	$293.00	
30-34	73	$2.93	$214.00	
35-39	41	$3.60	$147.67	
40+	13	$2.94	$38.24	

Age Ranges	
<10	$4.54
10-14	$3.76
15-19	$3.86
20-24	$4.32
25-29	$3.81
30-34	$4.12
35-39	$4.76
40+	$3.19

[8]:
```
top_spender_df1 = purchase_df.groupby('SN').sum()
total_purchase_value = top_spender_df1["Price"]
top_spender_df2 = purchase_df.groupby('SN').count()
purchase_count = top_spender_df2['Price']
average_purchase_price = total_purchase_value/purchase_count
top_spender_df = pd.DataFrame({'Purchase Count':purchase_count,
                               'Average Purchase Price': average_purchase_price,
                               'Total Purchase Value': total_purchase_value})
top_spender_df = top_spender_df.sort_values('Total Purchase Value', ascending =␣
↪False)
top_spender_df['Average Purchase Price'] = top_spender_df['Average Purchase␣
↪Price'].map("${:.2f}".format)
top_spender_df['Total Purchase Value'] = top_spender_df['Total Purchase Value'].
↪map("${:.2f}".format)
top_spender_df.head()
```

[8]:

SN	Purchase Count	Average Purchase Price	Total Purchase Value
Lisosia93	5	$3.79	$18.96
Idastidru52	4	$3.86	$15.45
Chamjask73	3	$4.61	$13.83
Iral74	4	$3.40	$13.62
Iskadarya95	3	$4.37	$13.10

[9]:
```
popular_item_df = purchase_df[['Item ID', 'Item Name','Price']]
popular_item_df = popular_item_df.groupby(['Item ID','Item Name'])
purchase_count = popular_item_df['Item ID'].count()
total_purchase_value = popular_item_df ['Price'].sum()
item_price = total_purchase_value/purchase_count
popular_df = pd.DataFrame ({'Purchase Count':purchase_count, 'Item Price' :␣
↪item_price,
                            'Total Purchase Value': total_purchase_value})
popular_df= popular_df.sort_values('Purchase Count',ascending = False)
popular_df['Item Price'] = popular_df['Item Price'].map("${:.2f}".format)
popular_df['Total Purchase Value'] = popular_df['Total Purchase Value'].map("${:
↪.2f}".format)
popular_df.head()
```

[9]:

Item ID	Item Name	Purchase Count	Item Price	Total Purchase Value
92	Final Critic	13	$4.61	$59.99
178	Oathbreaker, Last Hope of the Breaking Storm	12	$4.23	$50.76
145	Fiery Glass Crusader	9	$4.58	$41.22
132	Persuasion	9	$3.22	$28.99
108	Extraction, Quickblade Of Trembling Hands	9	$3.53	$31.77

[10]:
```
popular_item_df = purchase_df[['Item ID', 'Item Name','Price']]
popular_item_df = popular_item_df.groupby(['Item ID','Item Name'])
purchase_count = popular_item_df['Item ID'].count()
total_purchase_value = popular_item_df ['Price'].sum()
item_price = total_purchase_value/purchase_count
popular_df = pd.DataFrame ({'Purchase Count':purchase_count, 'Item Price' :␣
↪item_price,
                            'Total Purchase Value': total_purchase_value})
popular_df= popular_df.sort_values('Total Purchase Value',ascending = False)
popular_df['Item Price'] = popular_df['Item Price'].map("${:.2f}".format)
popular_df['Total Purchase Value'] = popular_df['Total Purchase Value'].map("${:
↪.2f}".format)
popular_df.head()
```

[10]:

Item ID	Item Name	Purchase Count	Item Price
92	Final Critic	13	
178	Oathbreaker, Last Hope of the Breaking Storm	12	
82	Nirvana	9	
145	Fiery Glass Crusader	9	
103	Singed Scalpel	8	

Item ID	Item Name	
92	Final Critic	$4.61
178	Oathbreaker, Last Hope of the Breaking Storm	$4.23
82	Nirvana	$4.90
145	Fiery Glass Crusader	$4.58
103	Singed Scalpel	$4.35

Item ID	Item Name	Total Purchase Value
92	Final Critic	$59.99
178	Oathbreaker, Last Hope of the Breaking Storm	$50.76
82	Nirvana	$44.10
145	Fiery Glass Crusader	$41.22
103	Singed Scalpel	$34.80

[ ]: