In [134]:
```
import pandas as pd
import numpy as np
from IPython.display import HTML, display
```

In [135]:
```
df = pd.read_json('purchase_data.json')
df.head()
```

Out[135]:

	Age	Gender	Item ID	Item Name	Price	SN
0	38	Male	165	Bone Crushing Silver Skewer	3.37	Aelalis34
1	21	Male	119	Stormbringer, Dark Blade of Ending Misery	2.32	Eolo46
2	34	Male	174	Primitive Blade	2.46	Assastnya25
3	21	Male	92	Final Critic	1.36	Pheusrical25
4	23	Male	63	Stormfury Mace	1.27	Aela59

In [136]:
```
total_players = len(df["SN"].value_counts())
data1 = [
    ["Total Number of Players: "],
    [total_players],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in data1))))
```

Total Number of Players:
573

In [137]:
```
unique_items = len(df["Item ID"].value_counts())
average_price = round(df["Price"].mean(), 2)
total_purchases = df.count()[0]
total_revenue = round(df["Price"].sum(), 2)

data2 = [
    ["Number of Unique Items: ", "Average Purchase Price: ", "Total Nu
mber of Purchases: ", "Total Revenue: "],
    [unique_items, average_price, total_purchases, total_revenue],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td
>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in dat
a2'))))
```

Number of Unique Items:	Average Purchase Price:	Total Number of Purchases:	Total Revenue:
183	2.93	780	2286.33

In [138]:
```
group_gender = df.groupby(by="Gender")
[female_count, male_count, other_count] = group_gender["SN"].nunique()
male_pct = male_count/(male_count + female_count + other_count)
female_pct = female_count/(male_count + female_count + other_count)
other_pct = other_count/(male_count + female_count + other_count)

data3 = [
    ["","Percentage","Count"],
    ["Male", str(round(male_pct*100)) + "%", male_count],
    ["Female", str(round(female_pct*100)) + "%", female_count],
    ["Other", str(round(other_pct*100)) + "%", other_count],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td
>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in dat
a3'))))
```

	Percentage	Count
Male	81.0%	465
Female	17.0%	100
Other	1.0%	8

In [143]:
```
[female_purchase, male_purchase, other_purchase] = group_gender["Age"]
.count()
[female_avg, male_avg, other_avg] = group_gender["Price"].mean()
[female_revenue, male_revenue, other_revenue] = group_gender["Price"].
sum()

data4 = [
    ["", "Purchase Count", "Average Purchase Price", "Total Purchase V
alue", "Normalized Totals"],
    ["Male", male_purchase, round(male_avg, 2), male_revenue, round(ma
le_revenue/male_count, 2)],
    ["Female", female_purchase, round(female_avg, 2), female_revenue,
round(female_revenue/female_count, 2)],
    ["Other", other_purchase, round(other_avg, 2), other_revenue, roun
d(other_revenue/other_count, 2)],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td
>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in dat
a4))))
```

	Purchase Count	Average Purchase Price	Total Purchase Value	Normalized Totals
Male	633	2.95	1867.68	4.02
Female	136	2.82	382.91	3.83
Other	11	3.25	35.74	4.47

In [152]:
```
bins = [0, 10, 20, 30, 40, 50]
labels = ["0-10", "11-20", "21-30", "31-40", "41-50"]
df["Age Group"] = pd.cut(df["Age"], bins, labels=labels)
group_age = df.groupby(by="Age Group")
[count_0, count_10, count_20, count_30, count_40] = group_age["SN"].nu
nique()
[purchase_0, purchase_10, purchase_20, purchase_30, purchase_40] = gro
up_age["SN"].count()
[avg_0, avg_10, avg_20, avg_30, avg_40] = group_age["Price"].mean()
[sum_0, sum_10, sum_20, sum_30, sum_40] = group_age["Price"].sum()

data5 = [
    ["Age", "Purchase Count", "Average Purchase Price", "Total Purchas
e Value", "Normalized Totals"],
    ["0-10", purchase_0, round(avg_0, 2), sum_0, round(sum_0/count_0,
2)],
    ["11-20", purchase_10, round(avg_10, 2), sum_10, round(sum_10/coun
t_10, 2)],
    ["21-30", purchase_20, round(avg_20, 2), sum_20, round(sum_20/coun
t_20, 2)],
    ["31-40", purchase_30, round(avg_30, 2), sum_30, round(sum_30/coun
t_30, 2)],
    ["41-50", purchase_40, round(avg_40, 2), sum_40, round(sum_40/coun
t_40, 2)],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td
>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in dat
a5))))
```

Age	Purchase Count	Average Purchase Price	Total Purchase Value	Normalized Totals
0-10	32	3.02	96.62	4.39
11-20	262	2.87	752.89	3.9
21-30	381	2.95	1122.43	3.92
31-40	102	3.0	305.75	4.43
41-50	3	2.88	8.64	2.88

In [213]:
```
group_top = df.groupby(by="SN")
spender5 = group_top["Price"].sum().sort_values(ascending=False)[0:5]
# spender5.index[0:5]
#group_top.count().loc[spender5.index[0]][0]

data6 = [
    ["SN", "Purchase Count", "Average Purchase Price", "Total Purchase
Value"],
    [spender5.index[0], group_top.count().loc[spender5.index[0]][0], r
ound(spender5[0]/group_top.count().loc[spender5.index[0]][0], 2), spen
der5[0]],
    [spender5.index[1], group_top.count().loc[spender5.index[1]][0], r
ound(spender5[1]/group_top.count().loc[spender5.index[1]][0], 2), spen
der5[1]],
    [spender5.index[2], group_top.count().loc[spender5.index[2]][0], r
ound(spender5[2]/group_top.count().loc[spender5.index[2]][0], 2), spen
der5[2]],
    [spender5.index[3], group_top.count().loc[spender5.index[3]][0], r
ound(spender5[3]/group_top.count().loc[spender5.index[4]][0], 2), spen
der5[3]],
    [spender5.index[4], group_top.count().loc[spender5.index[4]][0], r
ound(spender5[4]/group_top.count().loc[spender5.index[4]][0], 2), spen
der5[4]],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td
>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in dat
a6))))
```

SN	Purchase Count	Average Purchase Price	Total Purchase Value
Undirrala66	5	3.41	17.06
Saedue76	4	3.39	13.56
Mindimnya67	4	3.18	12.74
Haellysu29	3	4.24	12.73
Eoda93	3	3.86	11.58

In [241]:
```
group_pop = df.groupby(by="Item Name")
pop5 = group_pop["Item ID"].count().sort_values(ascending=False)[0:5]

data7 = [
    ["Item Name", "Purchase Count"],
    [pop5.index[0], group_pop.count().loc[pop5.index[0]][0]],
    [pop5.index[1], group_pop.count().loc[pop5.index[1]][0]],
    [pop5.index[2], group_pop.count().loc[pop5.index[2]][0]],
    [pop5.index[3], group_pop.count().loc[pop5.index[3]][0]],
    [pop5.index[4], group_pop.count().loc[pop5.index[4]][0]],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td
>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in dat
a7))))
```

Item Name	Purchase Count
Final Critic	14
Betrayal, Whisper of Grieving Widows	11
Arcane Gem	11
Stormcaller	10
Woeful Adamantite Claymore	9

In [240]:
```
sum5 = group_pop["Price"].sum().sort_values(ascending=False)[0:5]
data7 = [
    ["Item Name", "Total Purchase Value"],
    [sum5.index[0], sum5[0]],
    [sum5.index[1], sum5[1]],
    [sum5.index[2], sum5[2]],
    [sum5.index[3], sum5[3]],
    [sum5.index[4], sum5[4]],
    ]

display(HTML('<table><tr>{}</tr></table>'.format('</tr><tr>'.join('<td>{}</td>'.format('</td><td>'.join(str(_) for _ in row)) for row in data7))))
```

Item Name	Total Purchase Value
Final Critic	38.6
Retribution Axe	37.26
Stormcaller	34.65
Spectral Diamond Doomblade	29.75
Orenmir	29.7