

3mtj15wrc

April 7, 2025

```
[ ]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[ ]: df = pd.read_csv('Casino_Gaming_Data new.csv')
df.head()
```

```
[ ]: Fiscal Year      Month Ending      Licensee \
0      2021/22  06/30/2022 12:00:00 AM      Mohegan Tribe On-Reservation
1      2021/22  06/30/2022 12:00:00 AM      MPTN On-Reservation
2      2021/22  06/30/2022 12:00:00 AM  MPI Master Wagering License CT, LLC
3      2021/22  06/30/2022 12:00:00 AM      Mohegan Digital, LLC
4      2021/22  05/31/2022 12:00:00 AM      Mohegan Tribe On-Reservation
```

```
Wagers Patron Winnings Cancelled Wagers \
0      379023      335505      5
1      671843      685886      3
2  420056291  406814933  164727
3  254393582  246515856  11645
4      716607      681306      0
```

```
Online Casino Gaming Win/(Loss) \
0      43513
1     -14046
2    13076631
3    7866081
4     35301
```

```
Promotional Coupons or Credits Wagered (3) Promotional Deduction (4) \
0      6390      6390
1     29041      0
2   12372142  3269158
3   1397441  1397441
4     4130     4130
```

Total Gross Gaming Revenue Payment (5)

```

0          37123          6682
1         -14046           0
2        9807473       1765345
3       6468640       1164355
4         31171         5611

```

```
[ ]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 143 entries, 0 to 142
Data columns (total 11 columns):
#   Column                                          Non-Null Count  Dtype
---  -
0   Fiscal Year                                  143 non-null    object
1   Month Ending                                143 non-null    object
2   Licensee                                     143 non-null    object
3   Wagers                                       143 non-null    int64
4   Patron Winnings                             143 non-null    int64
5   Cancelled Wagers                           143 non-null    int64
6   Online Casino Gaming Win/(Loss)             143 non-null    int64
7   Promotional Coupons or Credits Wagered (3)  143 non-null    int64
8   Promotional Deduction (4)                   143 non-null    int64
9   Total Gross Gaming Revenue                  143 non-null    int64
10  Payment (5)                                 143 non-null    int64
dtypes: int64(8), object(3)
memory usage: 12.4+ KB

```

```
[ ]: df.describe()
```

```

[ ]:
      Wagers  Patron Winnings  Cancelled Wagers  \
count  1.430000e+02    1.430000e+02    1.430000e+02
mean   2.742207e+08    2.655218e+08    1.783408e+05
std    2.815530e+08    2.725166e+08    2.826004e+05
min    1.145280e+05    1.113650e+05    0.000000e+00
25%    1.407190e+06    1.357855e+06    5.000000e+00
50%    2.656540e+08    2.575127e+08    3.887000e+03
75%    5.352846e+08    5.156584e+08    2.913700e+05
max    7.838842e+08    7.573176e+08    1.133136e+06

      Online Casino Gaming Win/(Loss)  \
count                               1.430000e+02
mean                               8.520509e+06
std                                8.921912e+06
min                                -8.099000e+04
25%                                3.713950e+04
50%                                7.866081e+06
75%                                1.587461e+07

```

max 2.570443e+07

	Promotional Coupons or Credits Wagered (3)	Promotional Deduction (4) \
count	1.430000e+02	1.430000e+02
mean	2.909760e+06	1.470359e+06
std	5.194974e+06	1.512958e+06
min	0.000000e+00	0.000000e+00
25%	1.118100e+04	6.188500e+03
50%	1.065100e+06	1.021521e+06
75%	3.844558e+06	3.071548e+06
max	3.220656e+07	4.163625e+06

	Total Gross Gaming Revenue	Payment (5)
count	1.430000e+02	1.430000e+02
mean	7.050150e+06	1.269229e+06
std	7.483596e+06	1.346855e+06
min	-8.099000e+04	0.000000e+00
25%	3.172700e+04	5.711000e+03
50%	6.821088e+06	1.227796e+06
75%	1.304302e+07	2.347743e+06
max	2.281557e+07	4.106803e+06

```
[ ]: df.isnull().sum()
```

```
[ ]: Fiscal Year          0
      Month Ending        0
      Licensee            0
      Wagers              0
      Patron Winnings     0
      Cancelled Wagers    0
      Online Casino Gaming Win/(Loss)  0
      Promotional Coupons or Credits Wagered (3)  0
      Promotional Deduction (4)  0
      Total Gross Gaming Revenue  0
      Payment (5)         0
      dtype: int64
```

```
[ ]: df.tail()
```

```
[ ]:      Fiscal Year      Month Ending      Licensee \
138    2024/25  08/31/2024 12:00:00 AM      Mohegan Digital, LLC
139    2024/25  07/31/2024 12:00:00 AM      Mohegan Tribe On-Reservation
140    2024/25  07/31/2024 12:00:00 AM      MPTN On-Reservation
141    2024/25  07/31/2024 12:00:00 AM  MPI Master Wagering License CT, LLC
142    2024/25  07/31/2024 12:00:00 AM      Mohegan Digital, LLC

      Wagers  Patron Winnings  Cancelled Wagers \
```

138	552906853	532307920	468730
139	3543119	3410769	132
140	3199509	3072580	265
141	692210181	668971512	697914
142	535796500	516077716	561709

Online Casino Gaming Win/(Loss) \	
138	20130203
139	132219
140	126663
141	22540755
142	19157075

Promotional Coupons or Credits Wagered (3)		Promotional Deduction (4) \
138	4098131	3019530
139	20758	19833
140	15071	15071
141	4284383	3381113
142	4411186	2873561

Total Gross Gaming Revenue		Payment (5)
138	17110672	3079921
139	112386	20229
140	111592	20087
141	19159642	3448735
142	16283514	2931033

```
[ ]: df.dropna(inplace=True)
df.head(3)
```

[]:	Fiscal Year	Month Ending	Licensee \
0	2021/22	06/30/2022 12:00:00 AM	Mohegan Tribe On-Reservation
1	2021/22	06/30/2022 12:00:00 AM	MPTN On-Reservation
2	2021/22	06/30/2022 12:00:00 AM	MPI Master Wagering License CT, LLC

	Wagers	Patron Winnings	Cancelled Wagers \
0	379023	335505	5
1	671843	685886	3
2	420056291	406814933	164727

Online Casino Gaming Win/(Loss) \	
0	43513
1	-14046
2	13076631

Promotional Coupons or Credits Wagered (3)		Promotional Deduction (4) \
0	6390	6390

1	29041	0
2	12372142	3269158

	Total Gross Gaming Revenue	Payment (5)
0	37123	6682
1	-14046	0
2	9807473	1765345

```
[ ]: duplicates = df.duplicated()
      duplicates.sum()
```

```
[ ]: np.int64(0)
```

```
[ ]: df.dtypes
```

```
[ ]: Fiscal Year          object
      Month Ending        object
      Licensee            object
      Wagers              int64
      Patron Winnings     int64
      Cancelled Wagers    int64
      Online Casino Gaming Win/(Loss) int64
      Promotional Coupons or Credits Wagered (3) int64
      Promotional Deduction (4) int64
      Total Gross Gaming Revenue int64
      Payment (5)         int64
      dtype: object
```

```
[ ]: df.duplicated().sum()
```

```
[ ]: np.int64(0)
```

```
[ ]: numerical = df.select_dtypes(include=np.number)
      numerical.head()
```

```
[ ]:      Wagers  Patron Winnings  Cancelled Wagers  \
0      379023      335505      5
1      671843      685886      3
2  420056291  406814933  164727
3  254393582  246515856  11645
4      716607      681306      0
```

```
      Online Casino Gaming Win/(Loss)  \
0      43513
1     -14046
2    13076631
3    7866081
```

4 35301

	Promotional Coupons or Credits Wagered (3)	Promotional Deduction (4) \
0	6390	6390
1	29041	0
2	12372142	3269158
3	1397441	1397441
4	4130	4130

	Total Gross Gaming Revenue	Payment (5)
0	37123	6682
1	-14046	0
2	9807473	1765345
3	6468640	1164355
4	31171	5611

```
[ ]: df['Net revenue'] = df['Total Gross Gaming Revenue'] - df['Promotional_↵Deduction (4)']  
df.head()
```

[]:	Fiscal Year	Month Ending	Licensee \
0	2021/22	06/30/2022 12:00:00 AM	Mohegan Tribe On-Reservation
1	2021/22	06/30/2022 12:00:00 AM	MPTN On-Reservation
2	2021/22	06/30/2022 12:00:00 AM	MPI Master Wagering License CT, LLC
3	2021/22	06/30/2022 12:00:00 AM	Mohegan Digital, LLC
4	2021/22	05/31/2022 12:00:00 AM	Mohegan Tribe On-Reservation

	Wagers	Patron Winnings	Cancelled Wagers \
0	379023	335505	5
1	671843	685886	3
2	420056291	406814933	164727
3	254393582	246515856	11645
4	716607	681306	0

	Online Casino Gaming Win/(Loss) \
0	43513
1	-14046
2	13076631
3	7866081
4	35301

	Promotional Coupons or Credits Wagered (3)	Promotional Deduction (4) \
0	6390	6390
1	29041	0
2	12372142	3269158
3	1397441	1397441
4	4130	4130

	Total Gross Gaming Revenue	Payment (5)	Net revenue
0	37123	6682	30733
1	-14046	0	-14046
2	9807473	1765345	6538315
3	6468640	1164355	5071199
4	31171	5611	27041

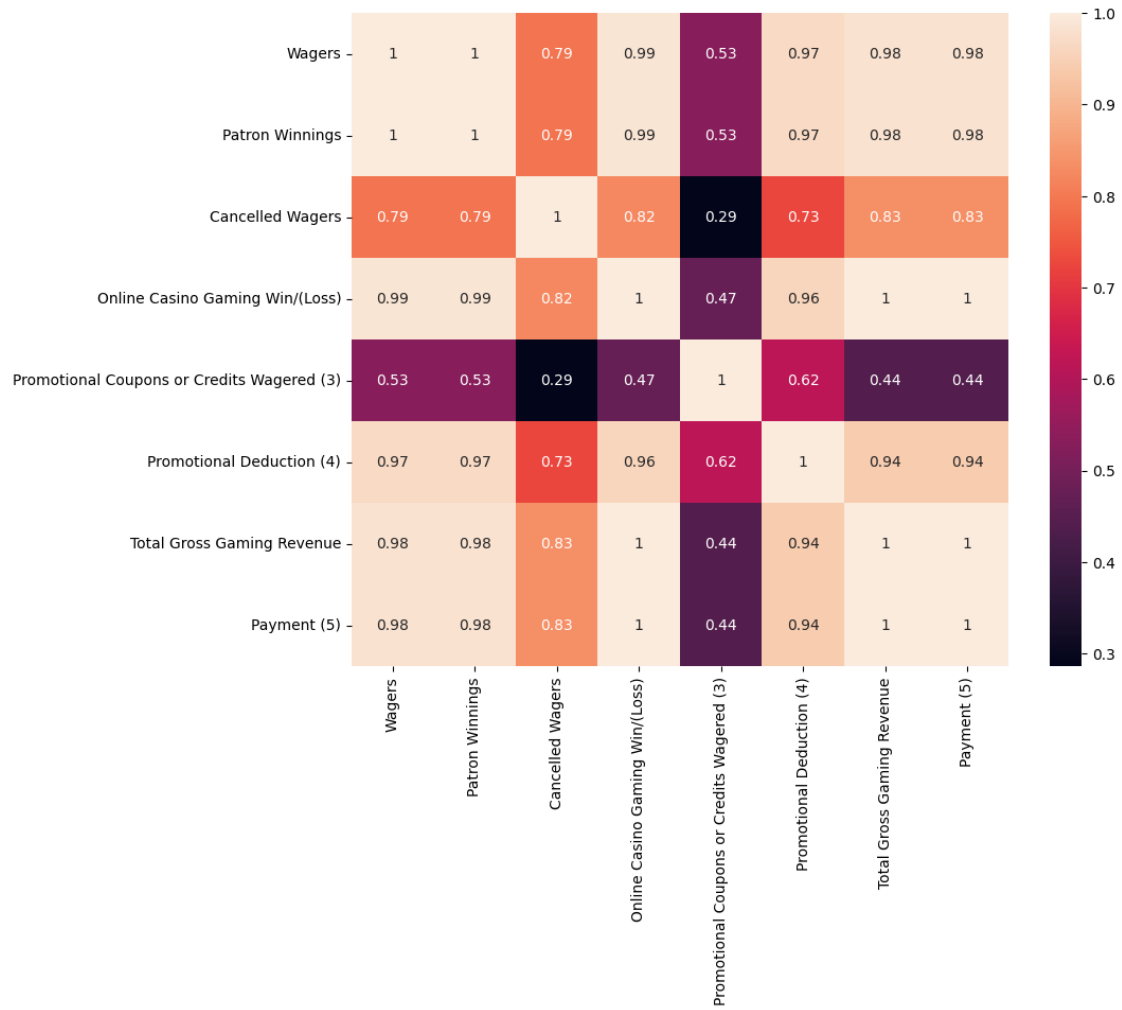
```
[ ]: # detect outliers
Q1 = numerical.quantile(0.25)
Q3 = numerical.quantile(0.75)
IQR = Q3 - Q1

lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

outliers = ((numerical < lower_bound) | (numerical > upper_bound)).any(axis=1)
outliers.sum()
```

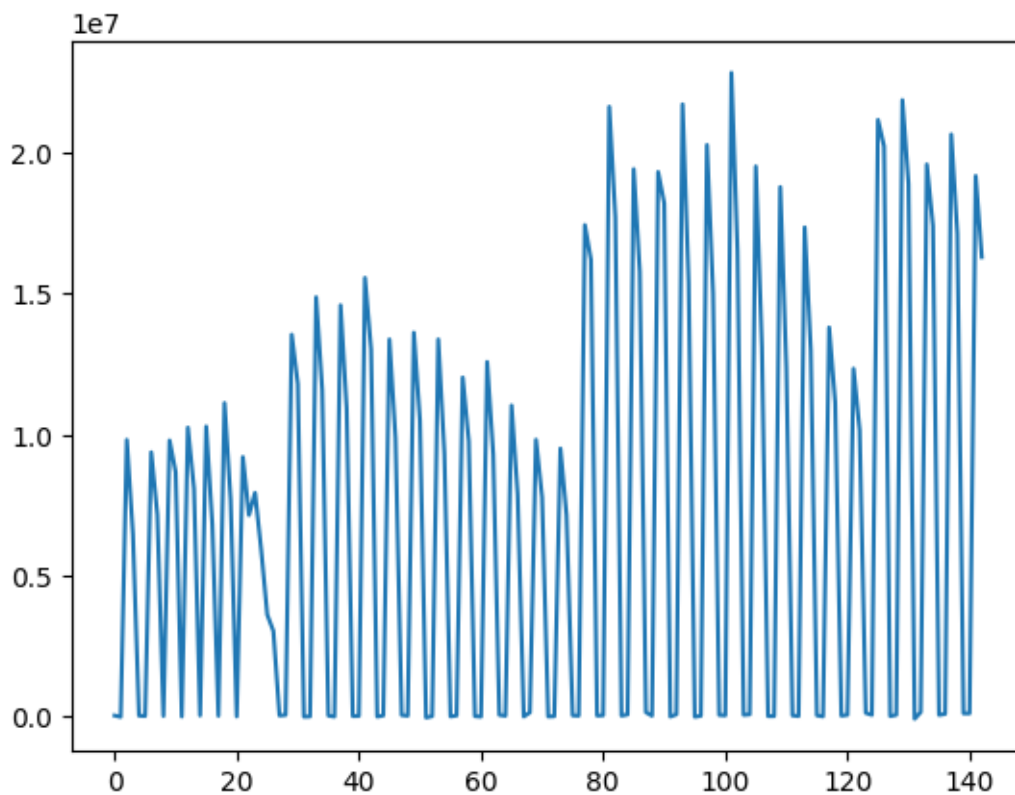
```
[ ]: np.int64(23)
```

```
[ ]: plt.figure(figsize=(10,8))
sns.heatmap(numerical.corr(), annot=True)
plt.show()
```



```
[ ]: df["Total Gross Gaming Revenue"].plot(kind='line')
```

```
[ ]: <Axes: >
```

```
[ ]: tggr = df['Total Gross Gaming Revenue']
      tggr
```

```
[ ]: 0      37123
      1     -14046
      2    9807473
      3    6468640
      4     31171
      ...
     138   17110672
     139    112386
     140    111592
     141   19159642
     142   16283514
      Name: Total Gross Gaming Revenue, Length: 143, dtype: int64
```

```
[ ]: df.columns
```

```
[ ]: Index(['Fiscal Year', 'Month Ending', 'Licensee', 'Wagers', 'Patron Winnings',
           'Cancelled Wagers', 'Online Casino Gaming Win/(Loss)',
           'Promotional Coupons or Credits Wagered (3)'],
```

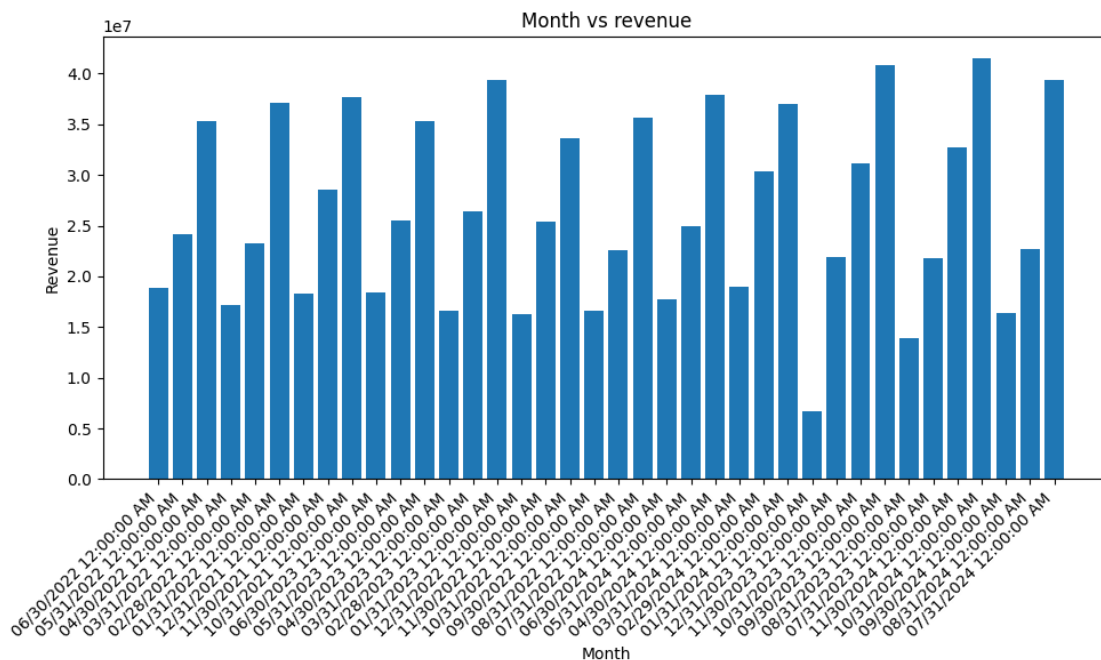
```
'Promotional Deduction (4)', 'Total Gross Gaming Revenue',
'Payment (5)', 'Net revenue'],
dtype='object')
```

```
[ ]: x = df['Month Ending']
x = pd.to_datetime(x)
```

<ipython-input-21-f7cf3e2d04d4>:2: UserWarning: Could not infer format, so each element will be parsed individually, falling back to `dateutil`. To ensure parsing is consistent and as-expected, please specify a format.

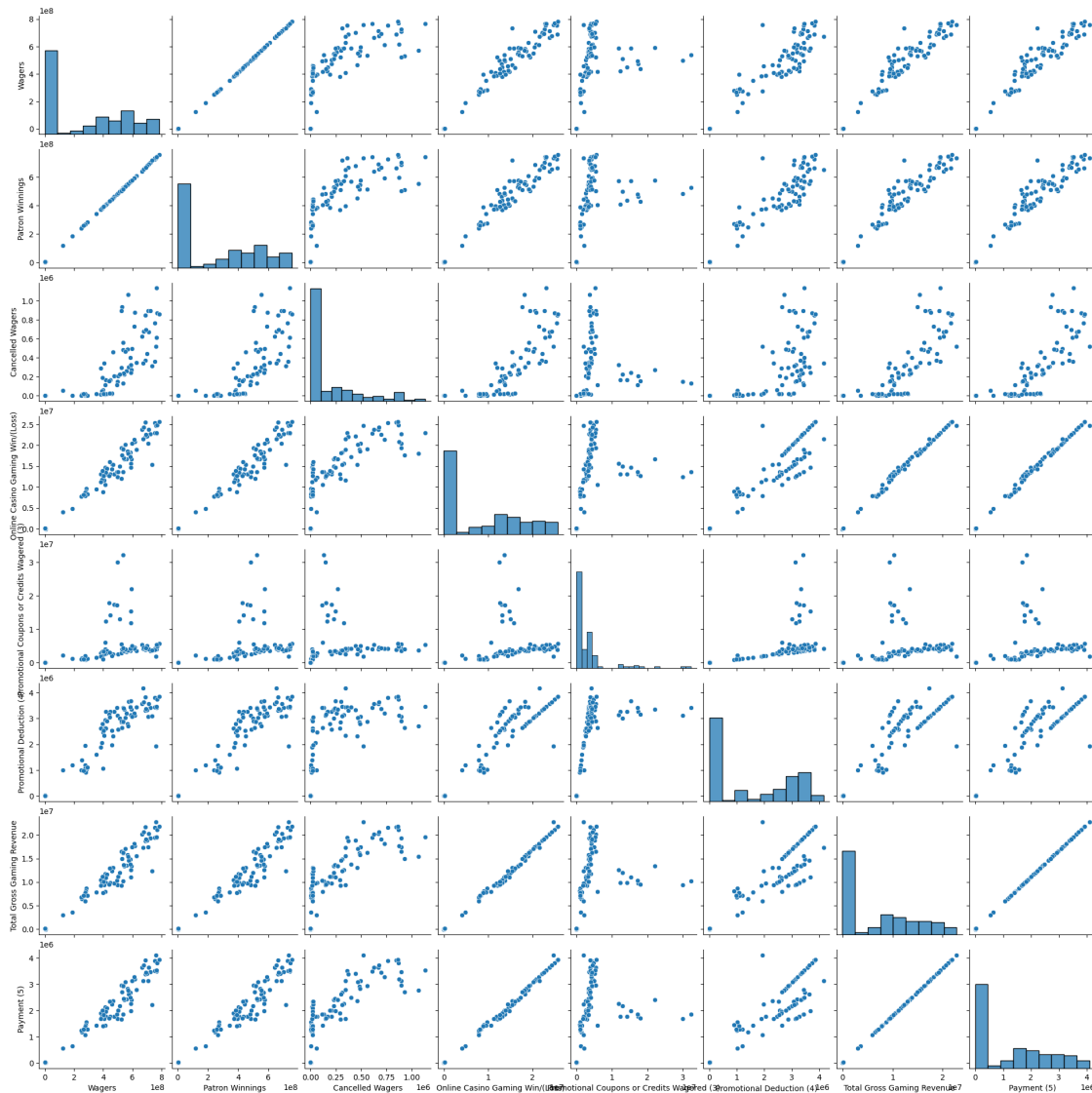
```
x = pd.to_datetime(x)
```

```
[ ]: trend = df.groupby('Month Ending')['Total Gross Gaming Revenue'].sum().mean()
plt.figure(figsize=(10,6))
plt.bar(df['Month Ending'].unique(), df.groupby('Month Ending')['Total Gross_
↪Gaming Revenue'].sum())
plt.xlabel("Month")
plt.ylabel("Revenue")
plt.title("Month vs revenue")
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```



```
[ ]: #sns.pairplot(numerical, diag_kind = "kde" )
plt.show()
```

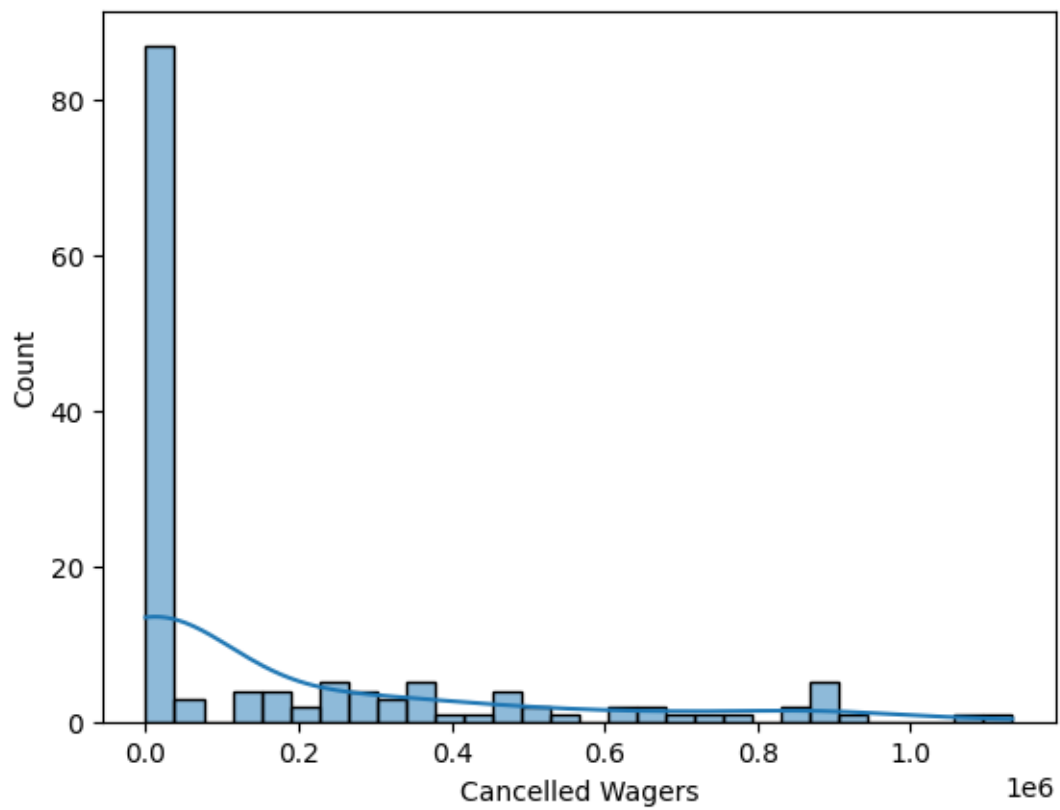
```
[ ]: sns.pairplot(numerical)
plt.show()
```



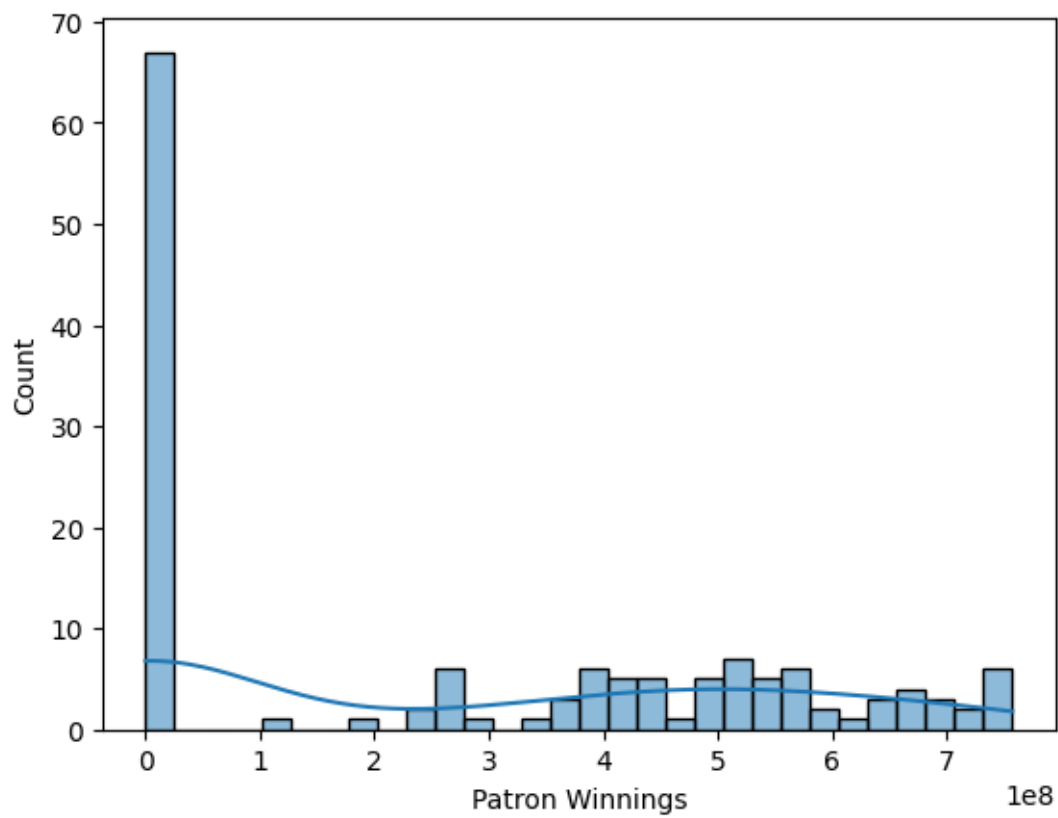
```
[ ]: numerical.columns
```

```
[ ]: Index(['Wagers', 'Patron Winnings', 'Cancelled Wagers',
            'Online Casino Gaming Win/(Loss)',
            'Promotional Coupons or Credits Wagered (3)',
            'Promotional Deduction (4)', 'Total Gross Gaming Revenue',
            'Payment (5)'],
          dtype='object')
```

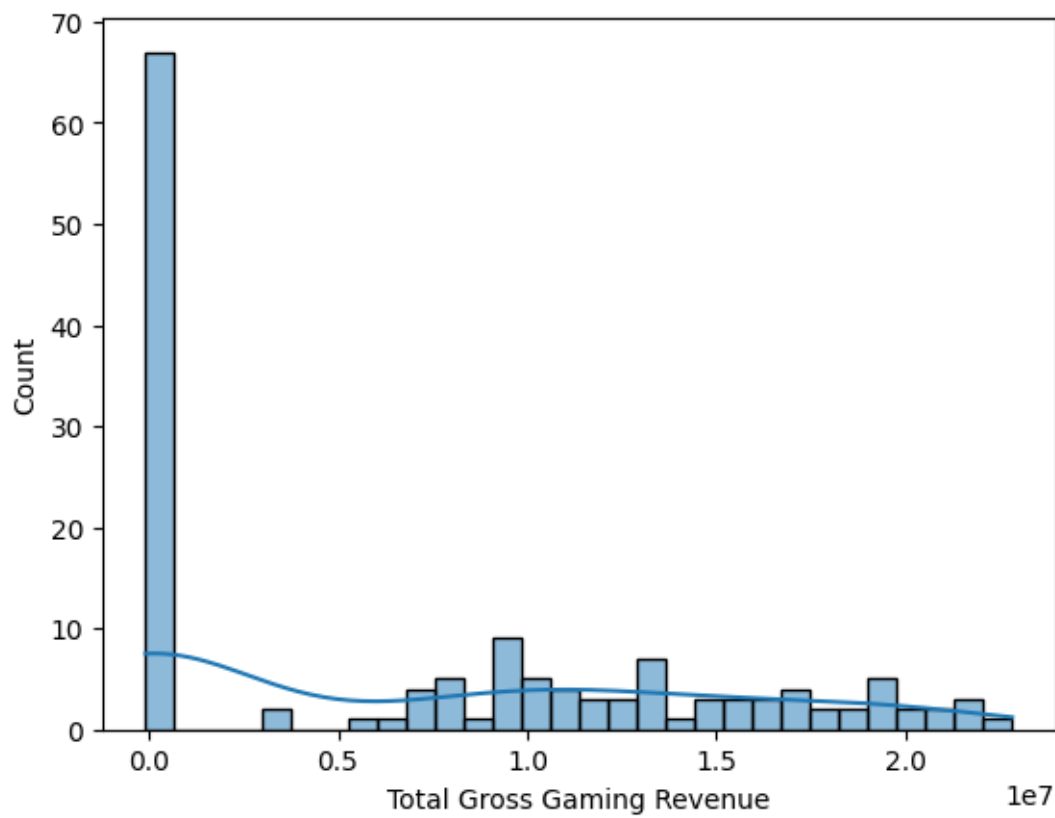
```
[ ]: sns.histplot(numerical['Cancelled Wagers'], bins = 30, kde = True)  
plt.show()
```



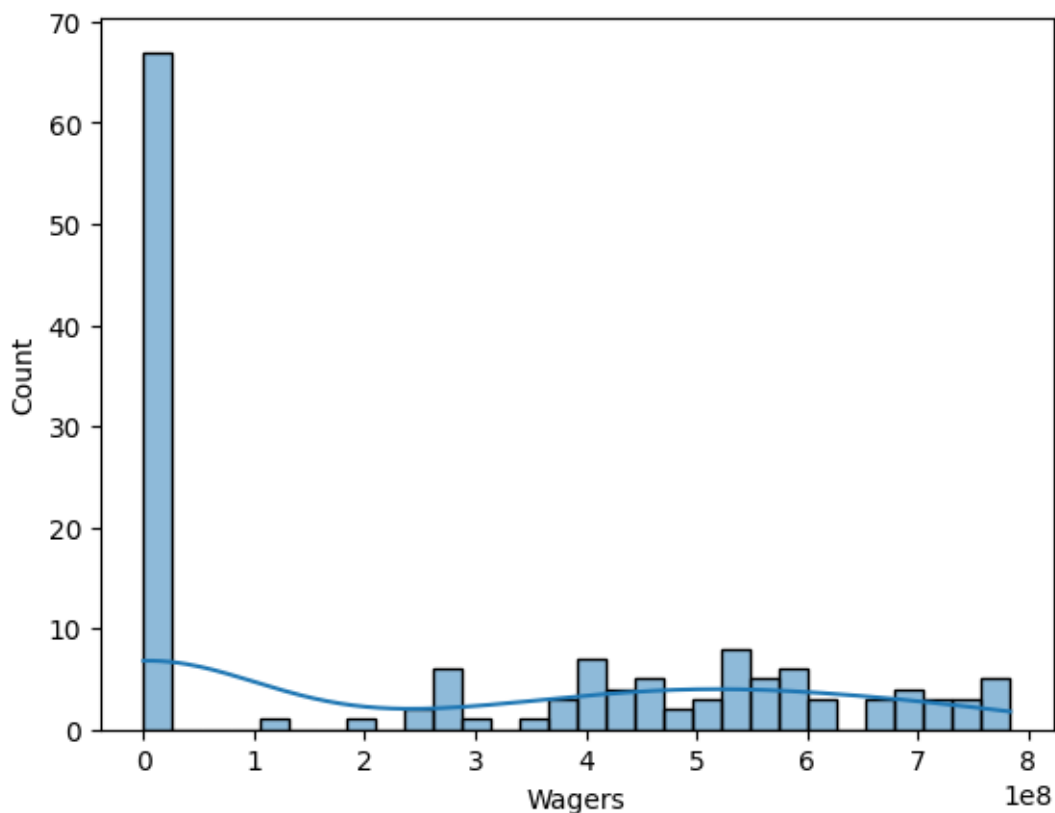
```
[ ]: sns.histplot(numerical['Patron Winnings'], bins = 30, kde = True)  
plt.show()
```



```
[ ]: sns.histplot(numerical['Total Gross Gaming Revenue'], bins = 30, kde = True)  
plt.show()
```



```
[ ]: sns.histplot(numerical['Wagers'], bins = 30, kde = True)  
plt.show()
```



```
[ ]: numerical.columns
```

```
[ ]: Index(['Wagers', 'Patron Winnings', 'Cancelled Wagers',
           'Online Casino Gaming Win/(Loss)',
           'Promotional Coupons or Credits Wagered (3)',
           'Promotional Deduction (4)', 'Total Gross Gaming Revenue',
           'Payment (5)'],
          dtype='object')
```

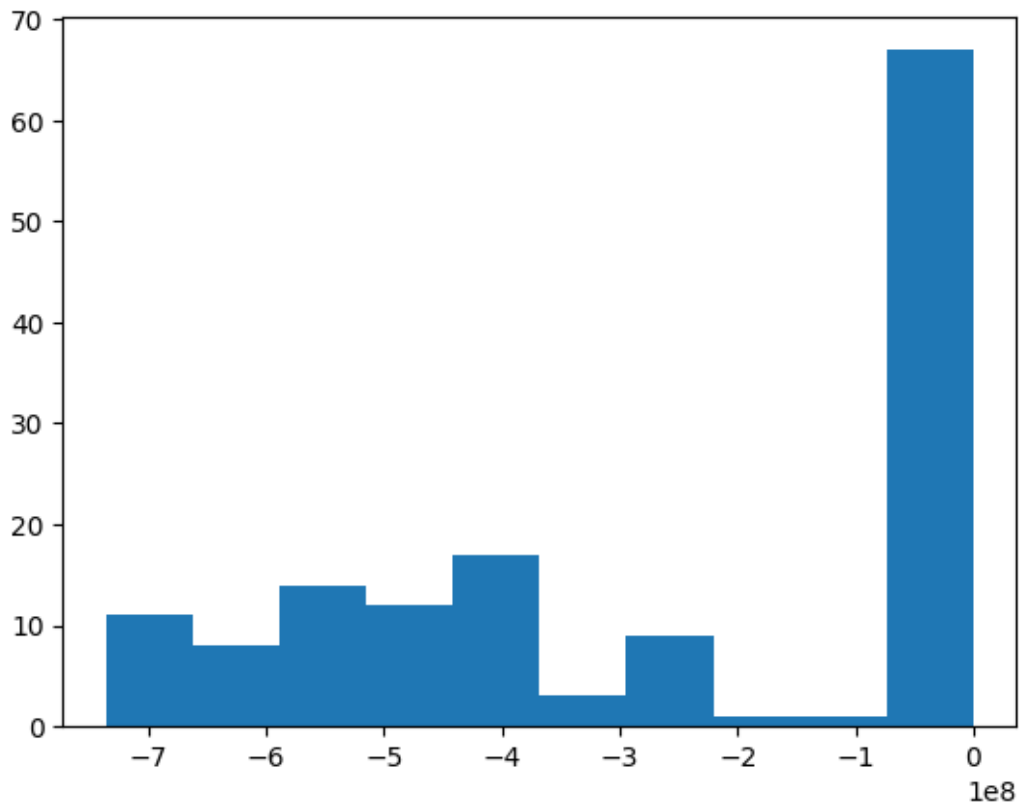
```
[ ]: numerical['Total Gross Gaming Revenue']
```

```
[ ]: 0      37123
     1     -14046
     2    9807473
     3   6468640
     4     31171
     ...
    138  17110672
    139   112386
    140   111592
    141  19159642
```

```
142    16283514
Name: Total Gross Gaming Revenue, Length: 143, dtype: int64
```

```
[ ]: df['Net revenue'] = df['Total Gross Gaming Revenue'] - df['Patron Winnings']
x = df['Net revenue']
```

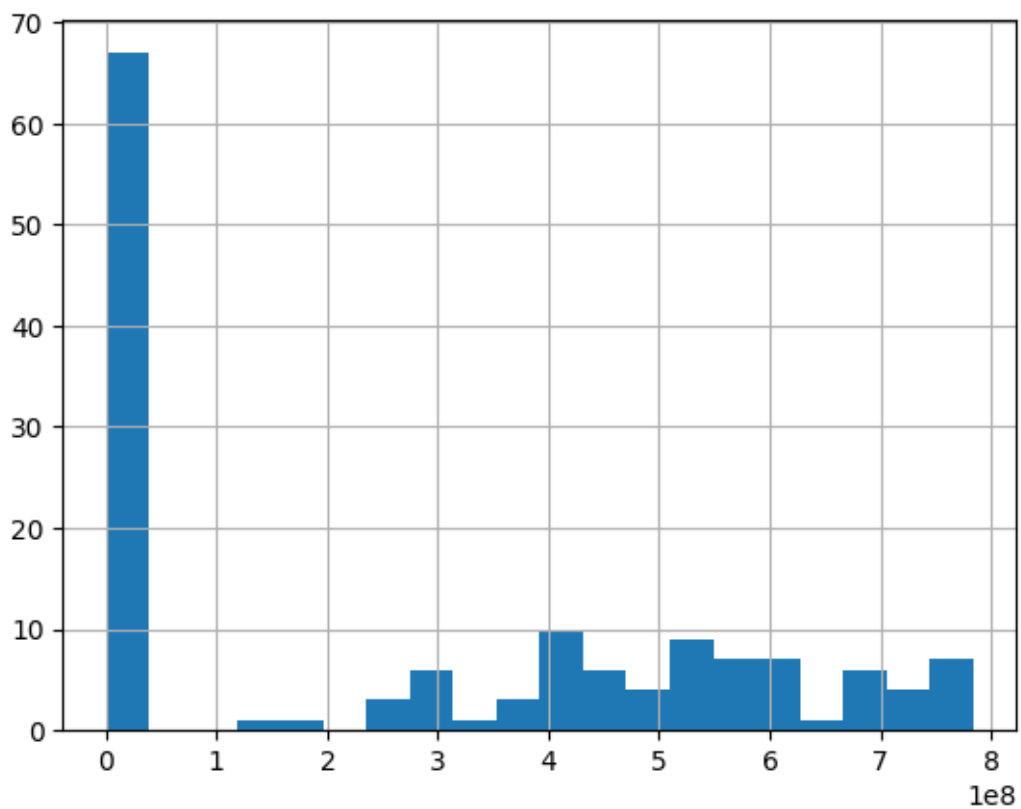
```
[ ]: plt.hist(x)
plt.show()
```



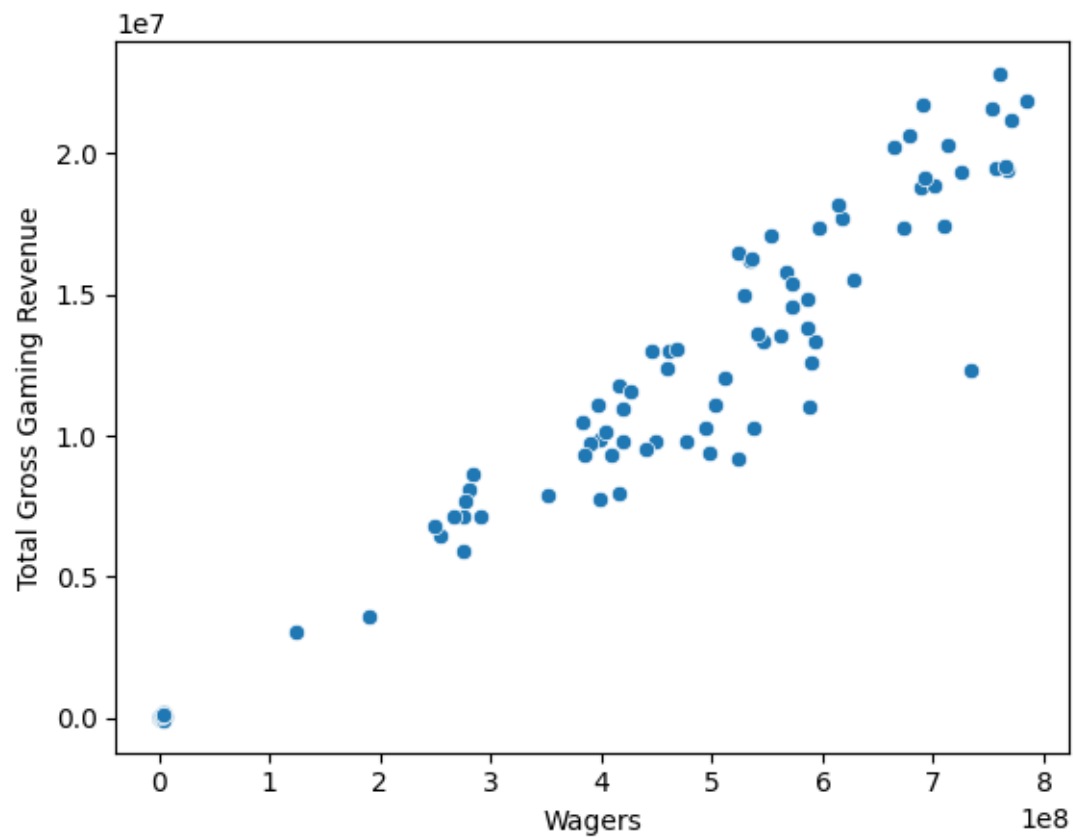
```
[ ]: df.columns
```

```
[ ]: Index(['Fiscal Year', 'Month Ending', 'Licensee', 'Wagers', 'Patron Winnings',
          'Cancelled Wagers', 'Online Casino Gaming Win/(Loss)',
          'Promotional Coupons or Credits Wagered (3)',
          'Promotional Deduction (4)', 'Total Gross Gaming Revenue',
          'Payment (5)', 'Net revenue'],
          dtype='object')
```

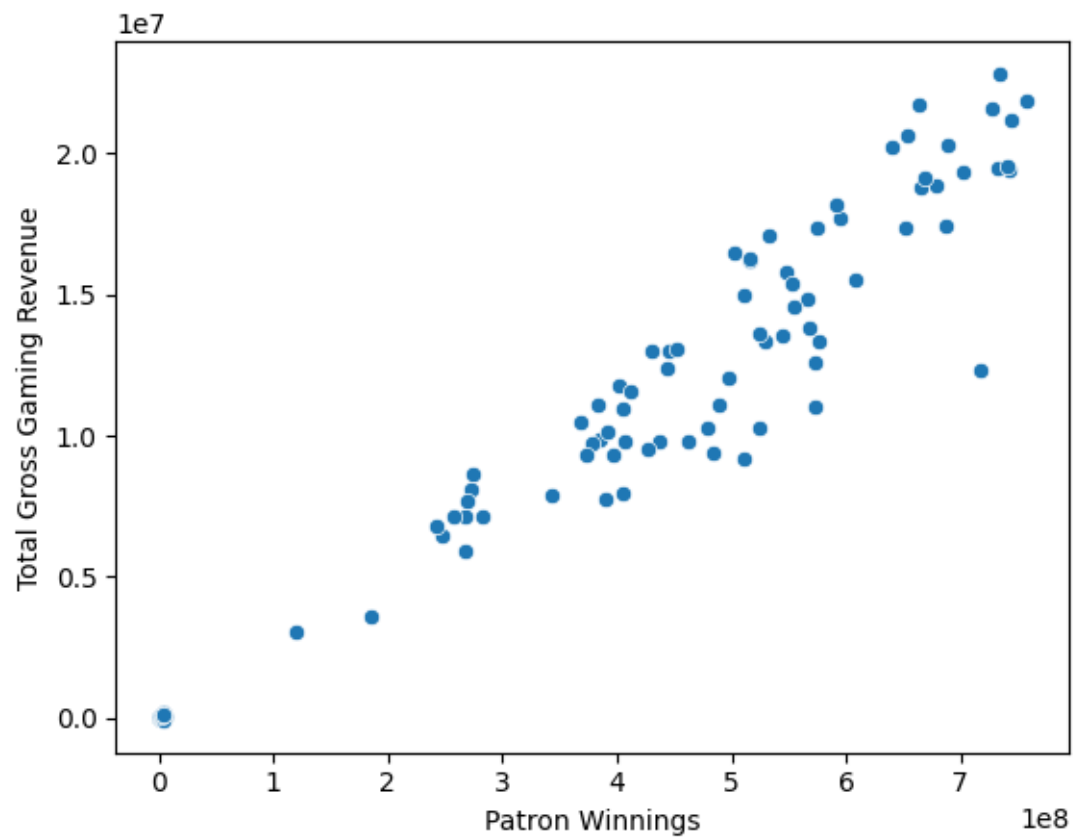
```
[ ]: df['Wagers'].hist(bins=20)
plt.show()
```

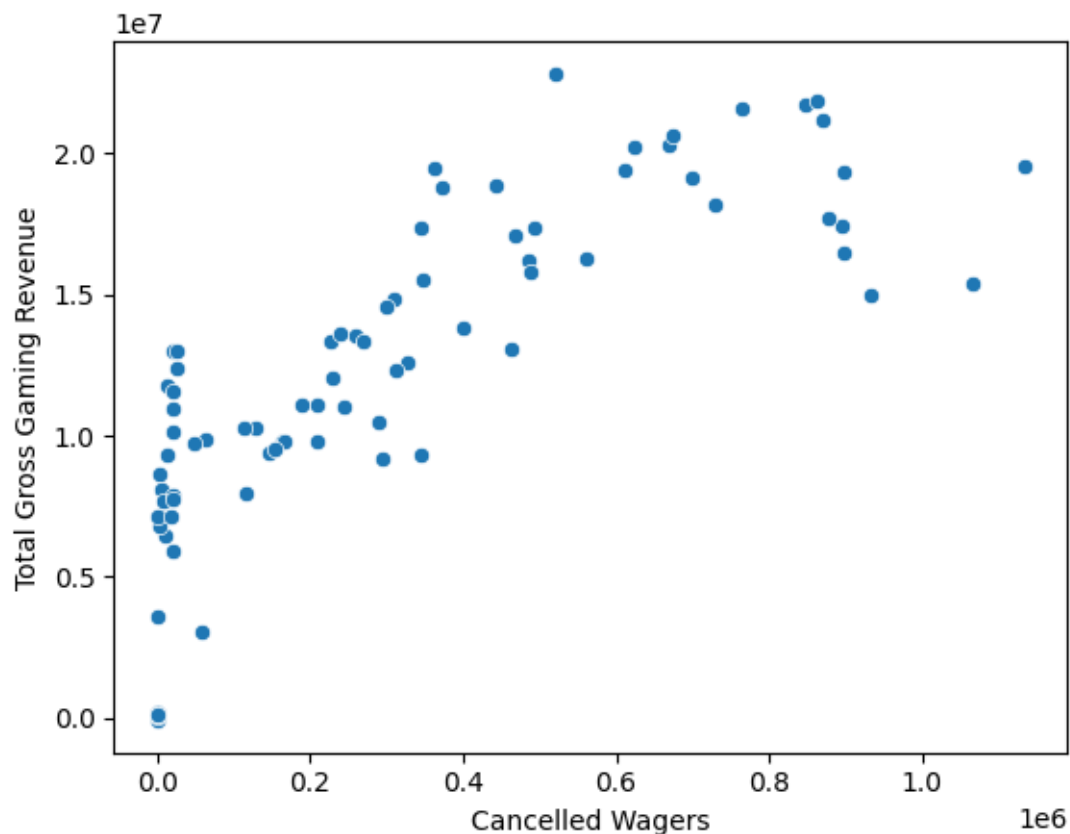
```
[ ]: sns.scatterplot(x='Wagers', y='Total Gross Gaming Revenue', data=df)
plt.show()
```



```
[ ]: #patron winnnings
sns.scatterplot(x='Patron Winnings', y='Total Gross Gaming Revenue', data=df)
plt.show()
```



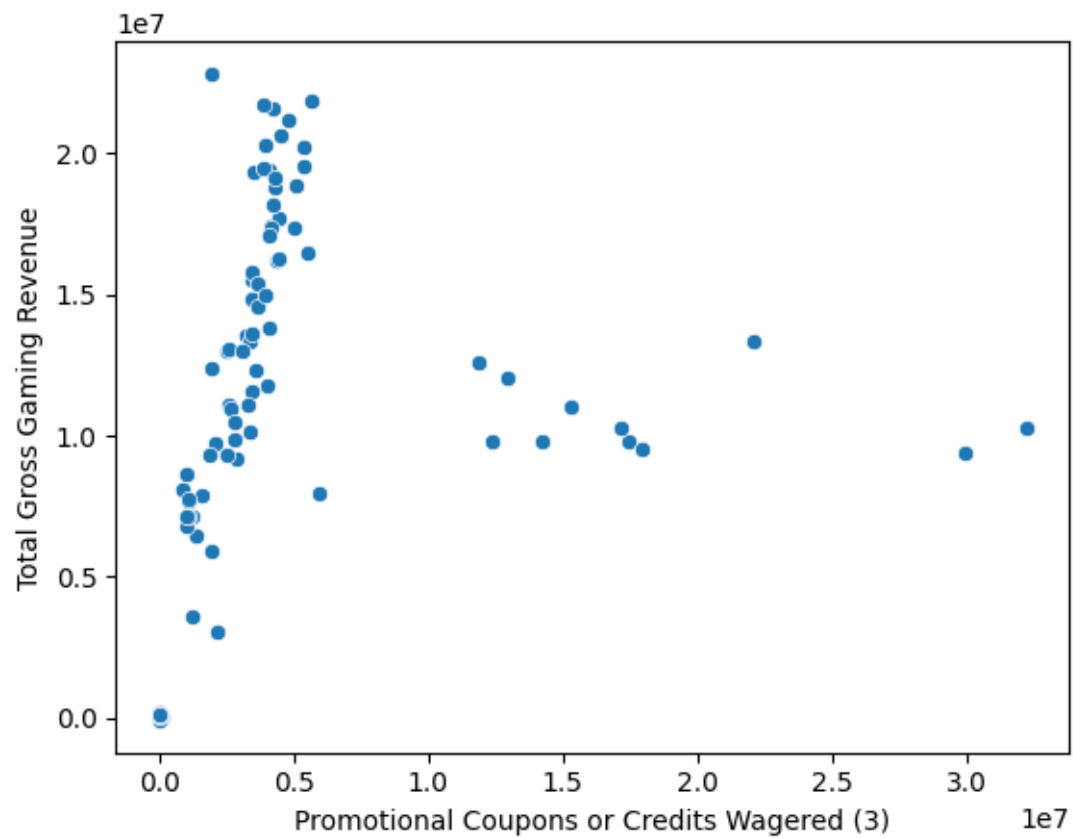
```
[ ]: #cancelled wagers  
sns.scatterplot(x='Cancelled Wagers', y='Total Gross Gaming Revenue', data=df)  
plt.show()
```



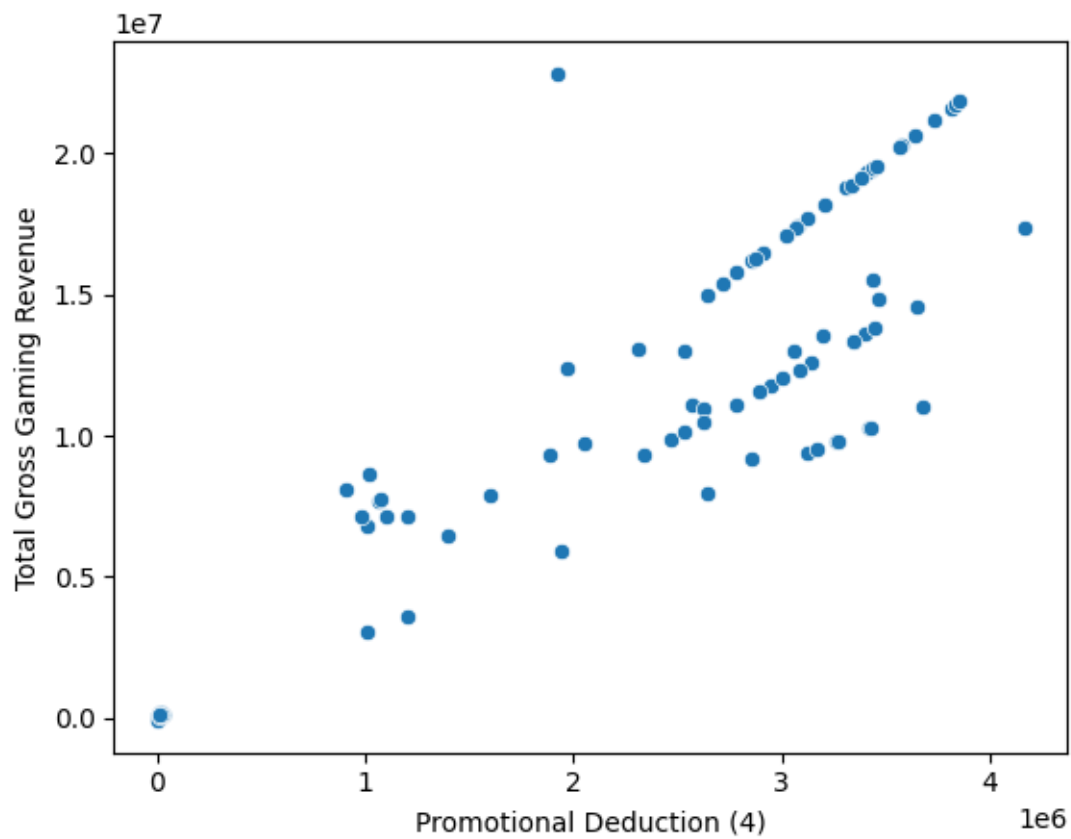
```
[ ]: df.columns
```

```
[ ]: Index(['Fiscal Year', 'Month Ending', 'Licensee', 'Wagers', 'Patron Winnings',
          'Cancelled Wagers', 'Online Casino Gaming Win/(Loss)',
          'Promotional Coupons or Credits Wagered (3)',
          'Promotional Deduction (4)', 'Total Gross Gaming Revenue',
          'Payment (5)', 'Net revenue'],
          dtype='object')
```

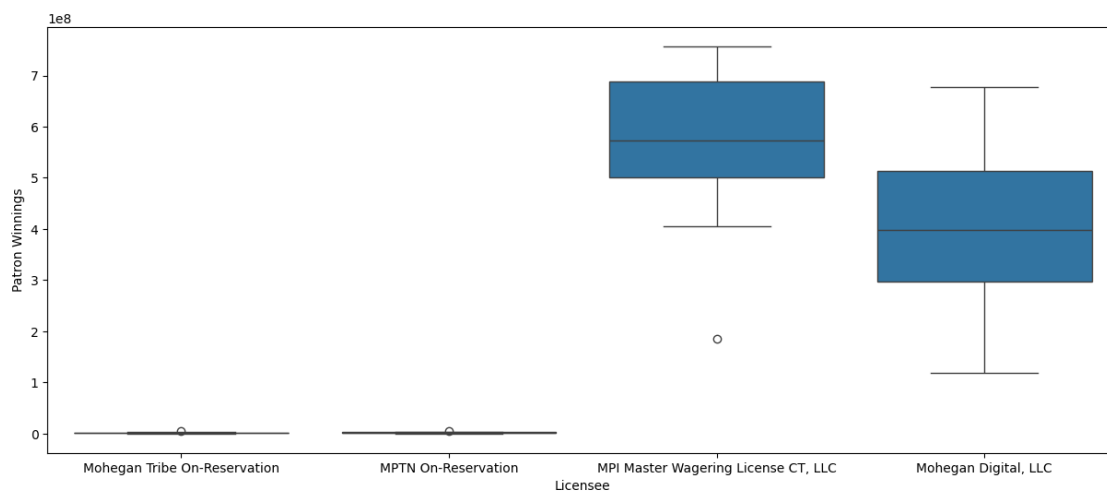
```
[ ]: sns.scatterplot(x='Promotional Coupons or Credits Wagered (3)', y='Total Gross_
    ↪Gaming Revenue', data=df)
plt.show()
```



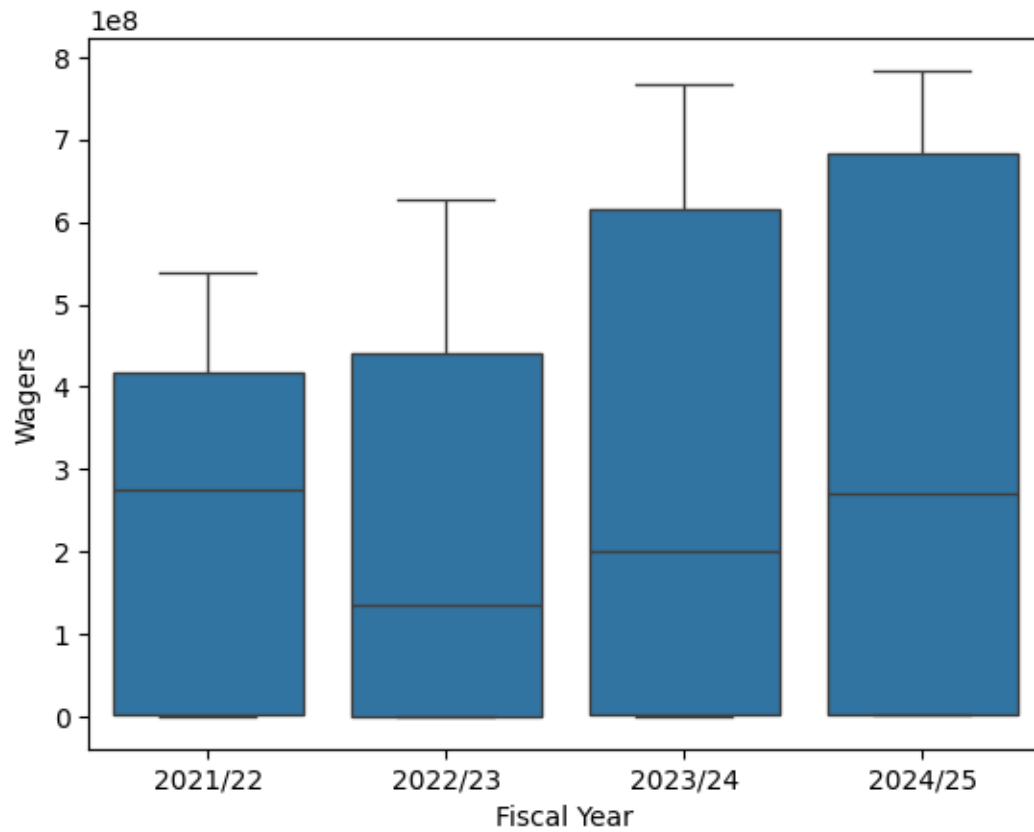
```
[ ]: sns.scatterplot(x='Promotional Deduction (4)', y='Total Gross Gaming Revenue',  
                    data=df)  
plt.show()
```



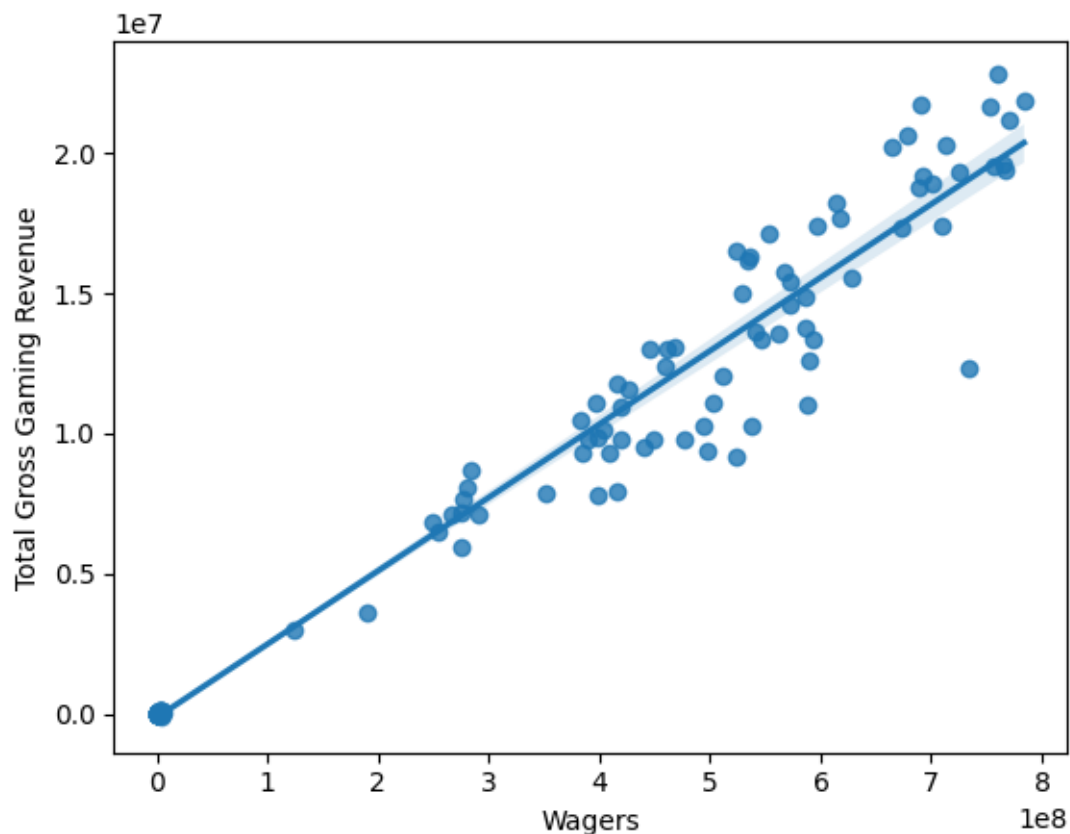
```
[ ]: #casino
plt.figure(figsize=(15, 6))
sns.boxplot(x='Licensee', y='Patron Winnings', data=df)
plt.show()
```



```
[ ]: sns.boxplot(x='Fiscal Year', y='Wagers', data=df)
plt.show()
```



```
[ ]: sns.regplot(x='Wagers', y='Total Gross Gaming Revenue', data=df)
plt.show()
```



```
[ ]: df['Net Winnings'] = df['Patron Winnings'] - df['Cancelled Wagers']
df.head()
```

```
[ ]: Fiscal Year      Month Ending      Licensee \
0      2021/22  06/30/2022  12:00:00 AM      Mohegan Tribe On-Reservation
1      2021/22  06/30/2022  12:00:00 AM      MPTN On-Reservation
2      2021/22  06/30/2022  12:00:00 AM      MPI Master Wagering License CT, LLC
3      2021/22  06/30/2022  12:00:00 AM      Mohegan Digital, LLC
4      2021/22  05/31/2022  12:00:00 AM      Mohegan Tribe On-Reservation

      Wagers  Patron Winnings  Cancelled Wagers \
0      379023      335505      5
1      671843      685886      3
2  420056291  406814933      164727
3  254393582  246515856      11645
4      716607      681306      0

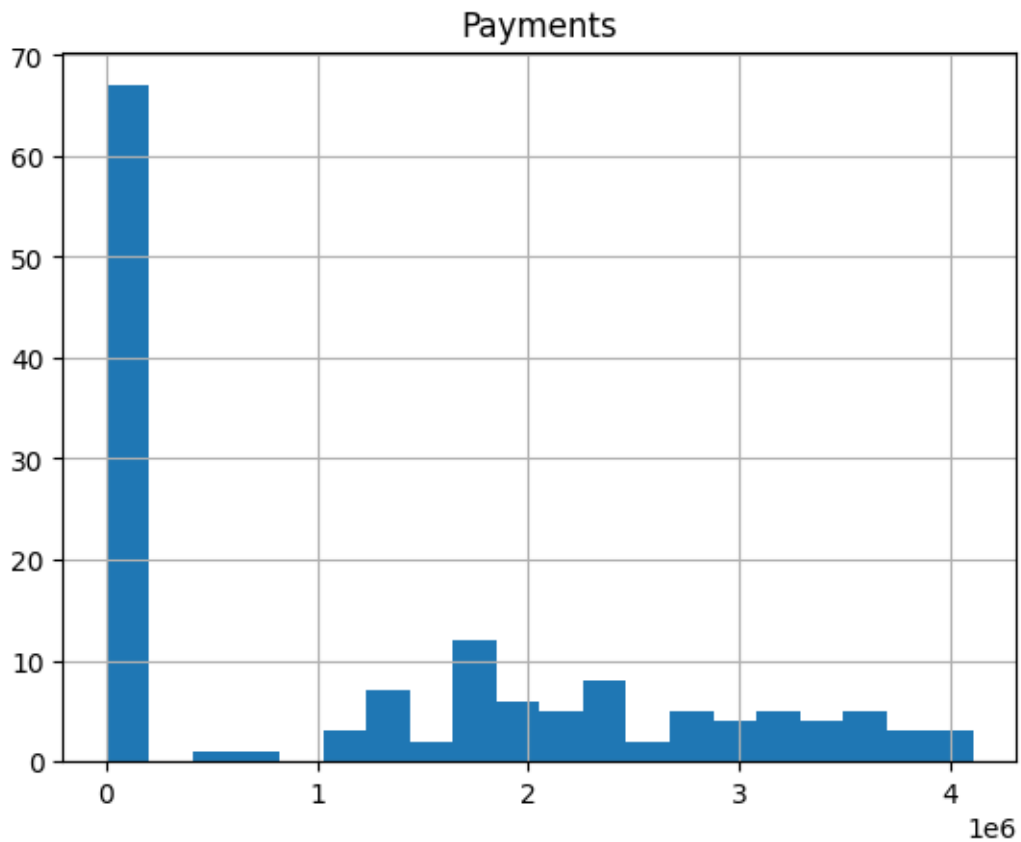
Online Casino Gaming Win/(Loss) \
0      43513
1     -14046
```


2	13076631
3	7866081
4	35301

	Promotional Coupons or Credits Wagered (3)	Promotional Deduction (4) \
0	6390	6390
1	29041	0
2	12372142	3269158
3	1397441	1397441
4	4130	4130

	Total Gross Gaming Revenue	Payment (5)	Net revenue	Net Winnings
0	37123	6682	-298382	335500
1	-14046	0	-699932	685883
2	9807473	1765345	-397007460	406650206
3	6468640	1164355	-240047216	246504211
4	31171	5611	-650135	681306

```
[ ]: df["Payment (5)"].hist(bins=20)
plt.title("Payments")
plt.show()
```



STATISTICAL ANALYSIS

```
[ ]: numerical.mean()
```

```
[ ]: Wagers                2.742207e+08
      Patron Winnings       2.655218e+08
      Cancelled Wagers      1.783408e+05
      Online Casino Gaming Win/(Loss)  8.520509e+06
      Promotional Coupons or Credits Wagered (3)  2.909760e+06
      Promotional Deduction (4)  1.470359e+06
      Total Gross Gaming Revenue  7.050150e+06
      Payment (5)           1.269229e+06
      dtype: float64
```

```
[ ]: numerical.median()
```

```
[ ]: Wagers                265653958.0
      Patron Winnings       257512736.0
      Cancelled Wagers           3887.0
      Online Casino Gaming Win/(Loss)  7866081.0
      Promotional Coupons or Credits Wagered (3)  1065100.0
      Promotional Deduction (4)  1021521.0
      Total Gross Gaming Revenue  6821088.0
      Payment (5)           1227796.0
      dtype: float64
```

```
[ ]: numerical.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 143 entries, 0 to 142
```

```
Data columns (total 8 columns):
```

#	Column	Non-Null Count	Dtype
0	Wagers	143 non-null	int64
1	Patron Winnings	143 non-null	int64
2	Cancelled Wagers	143 non-null	int64
3	Online Casino Gaming Win/(Loss)	143 non-null	int64
4	Promotional Coupons or Credits Wagered (3)	143 non-null	int64
5	Promotional Deduction (4)	143 non-null	int64
6	Total Gross Gaming Revenue	143 non-null	int64
7	Payment (5)	143 non-null	int64

```
dtypes: int64(8)
```

```
memory usage: 9.1 KB
```

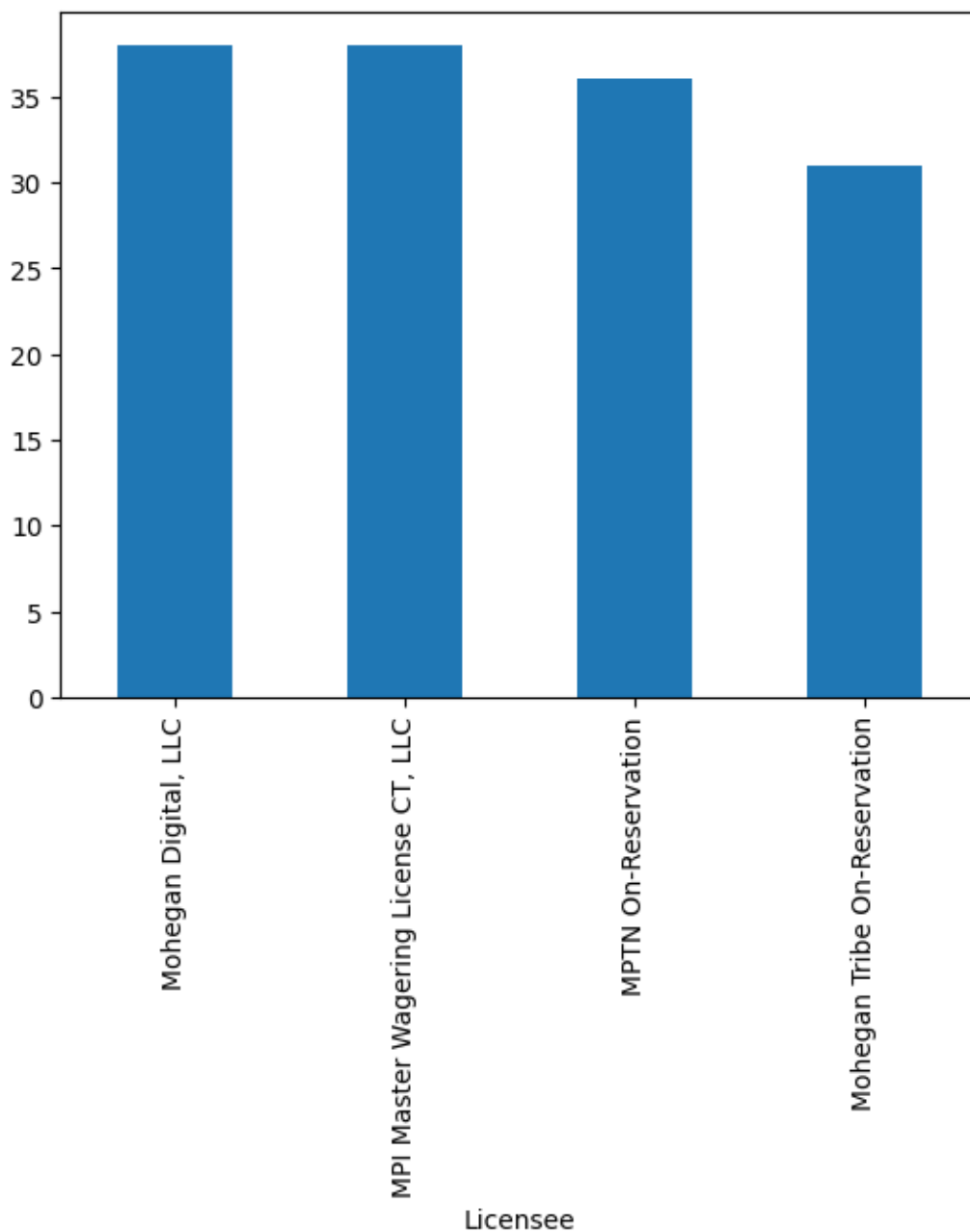
```
[ ]: numerical.std()
```

```
[ ]: Wagers 2.815530e+08
      Patron Winnings 2.725166e+08
      Cancelled Wagers 2.826004e+05
      Online Casino Gaming Win/(Loss) 8.921912e+06
      Promotional Coupons or Credits Wagered (3) 5.194974e+06
      Promotional Deduction (4) 1.512958e+06
      Total Gross Gaming Revenue 7.483596e+06
      Payment (5) 1.346855e+06
      dtype: float64
```

0.1 PROBLEM STATEMENTS AND INSIGHTS

```
[ ]: # revenue trend analysis
      df['Licensee'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Licensee'>
```



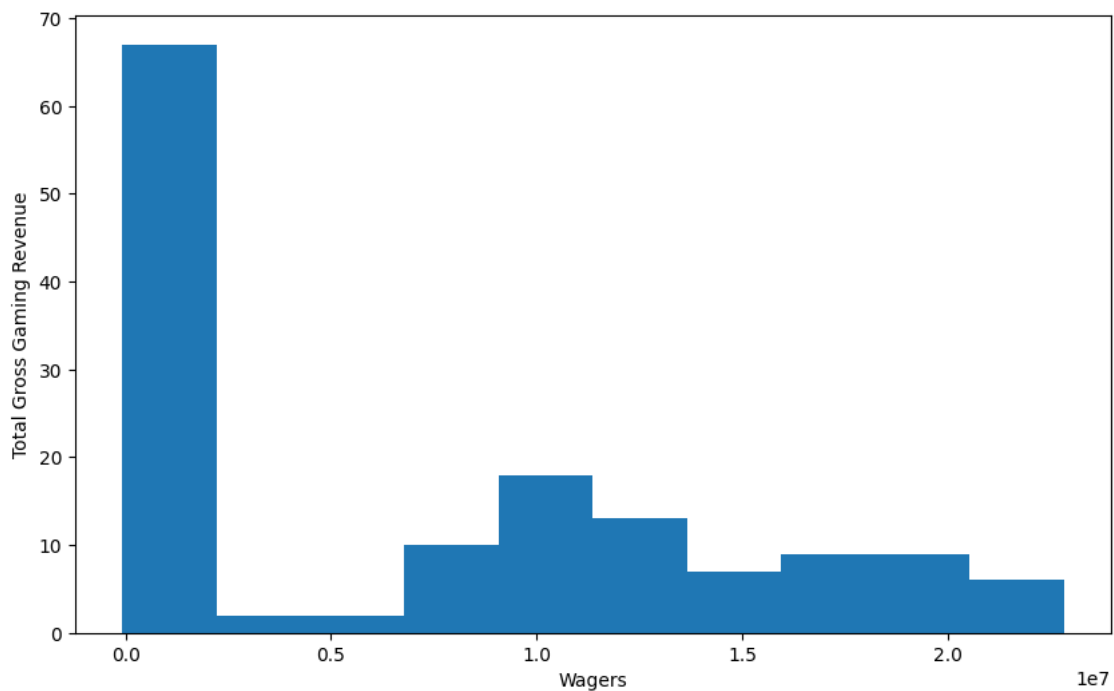
```
[ ]: casinos = df.groupby('Licensee')['Total Gross Gaming Revenue'].sum()
casinos
```

```
[ ]: Licensee
MPI Master Wagering License CT, LLC    561714816
MPTN On-Reservation                    1424731
Mohegan Digital, LLC                  444084162
Mohegan Tribe On-Reservation           947694
```

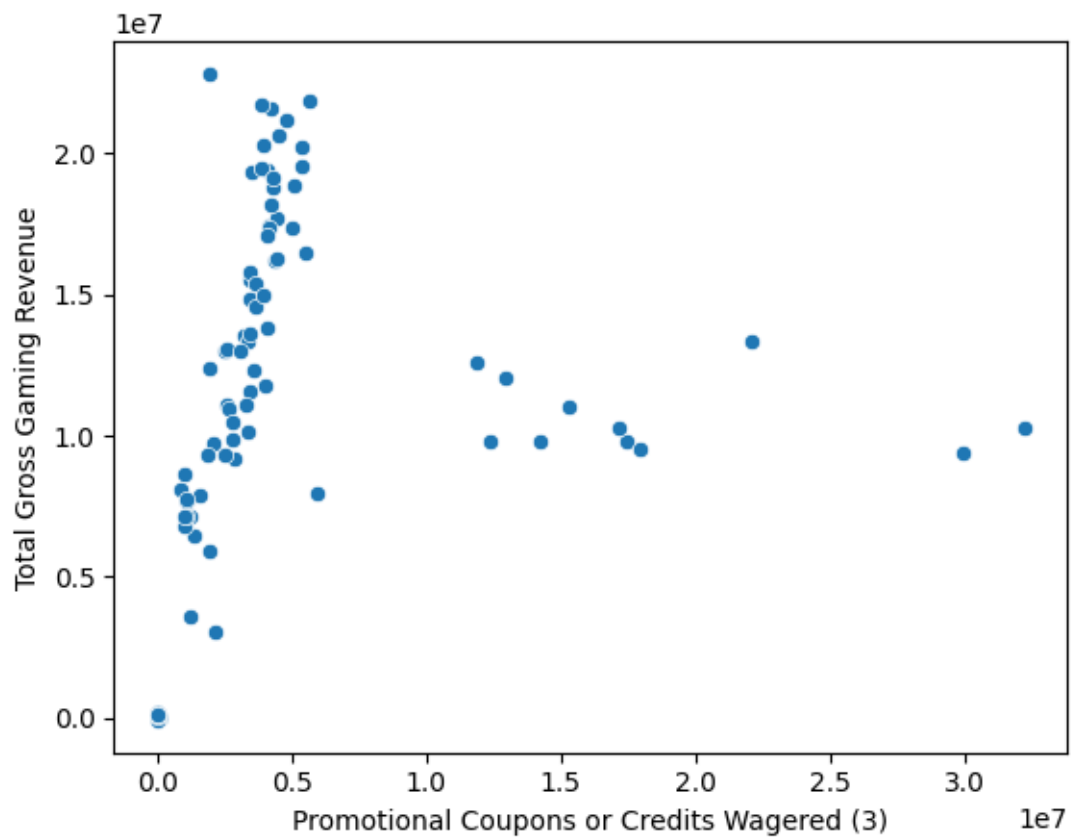
Name: Total Gross Gaming Revenue, dtype: int64

```
[ ]: wagers = df.groupby('Wagers')['Total Gross Gaming Revenue'].sum()
wagers

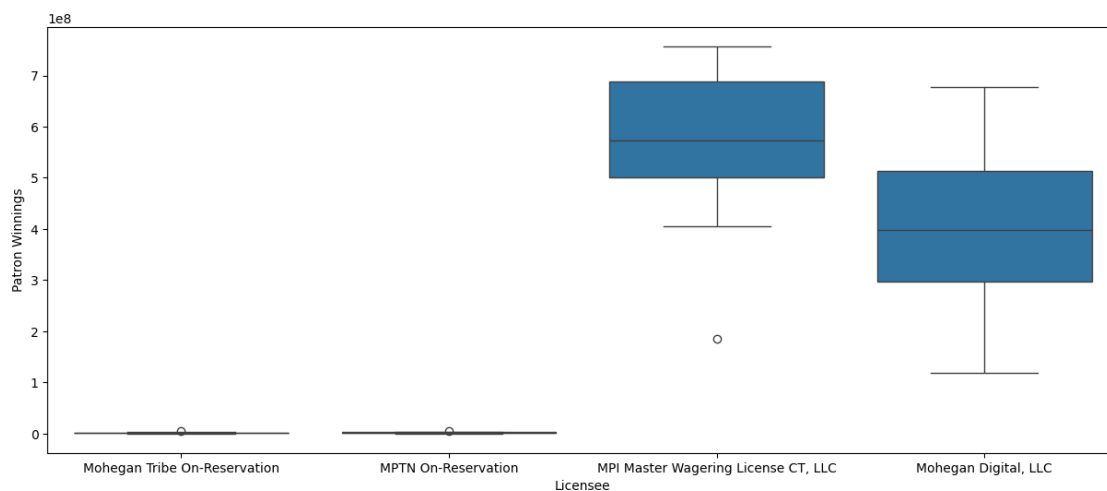
plt.figure(figsize=(10,6))
plt.hist(wagers)
plt.xlabel('Wagers')
plt.ylabel('Total Gross Gaming Revenue')
plt.show()
```



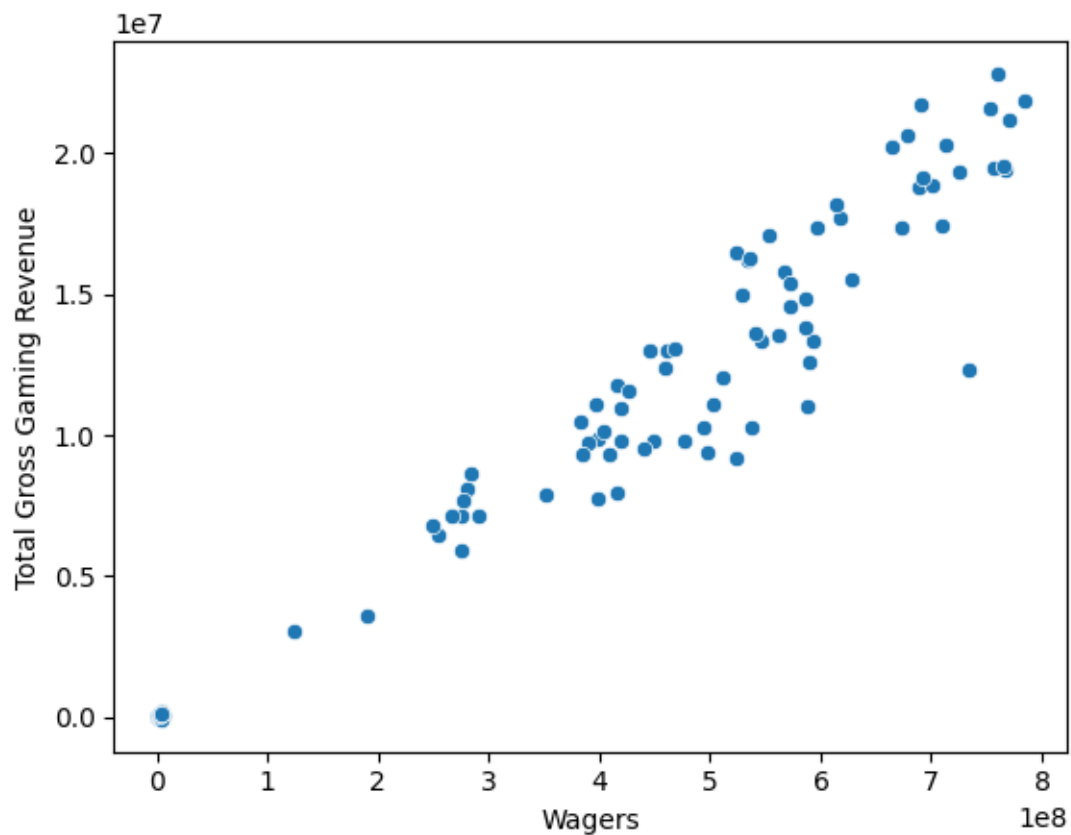
```
[ ]: # impact of promotions
sns.scatterplot(x='Promotional Coupons or Credits Wagered (3)', y='Total Gross Gaming Revenue', data=df)
plt.show()
```



```
[ ]: # casino performance comperasion
plt.figure(figsize=(15, 6))
sns.boxplot(x='Licensee', y='Patron Winnings', data=df)
plt.show()
```

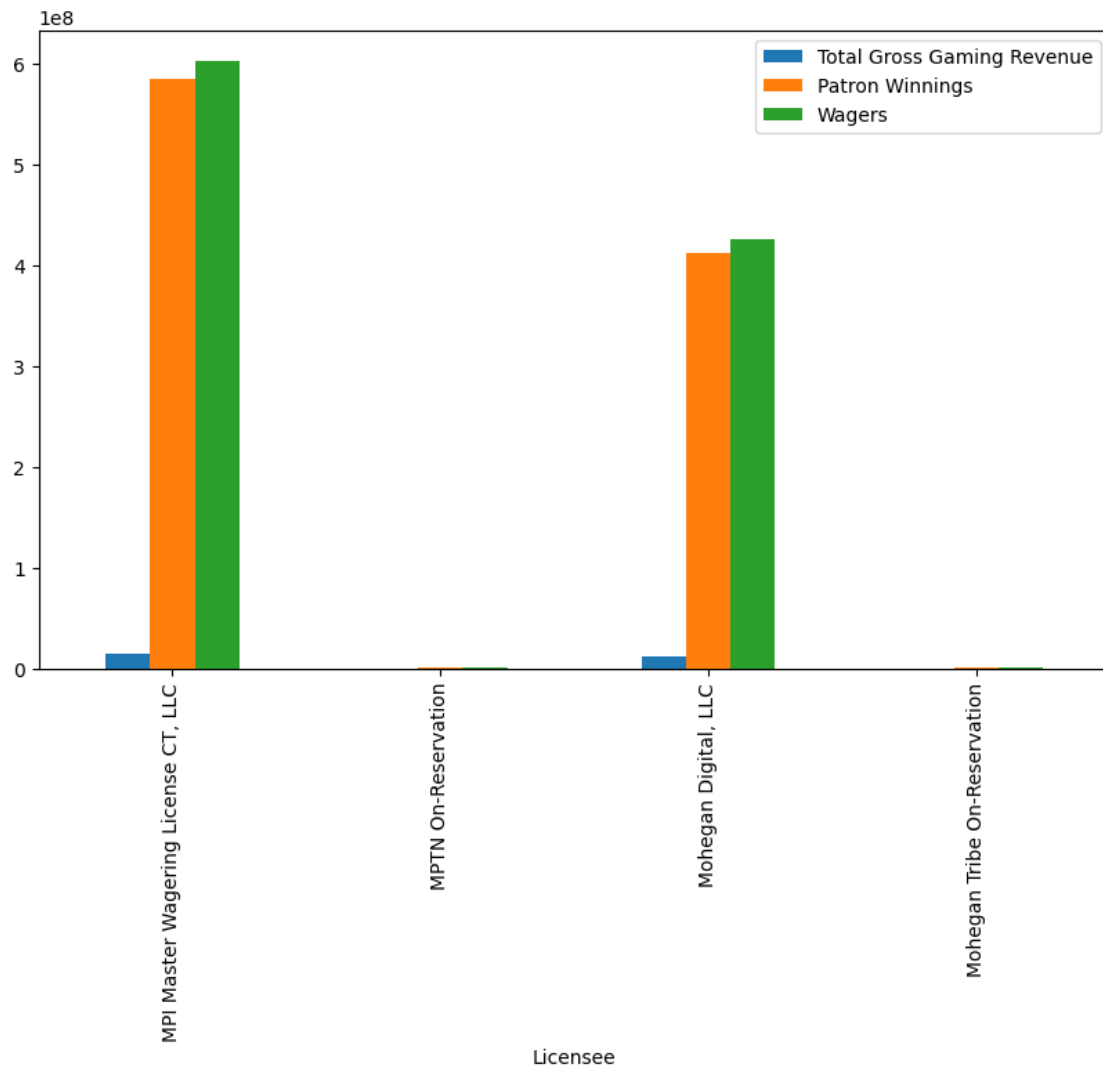


```
[ ]: # wagering payout patterns
sns.scatterplot(x='Wagers', y='Total Gross Gaming Revenue', data=df)
plt.show()
```



```
[ ]: df.groupby('Licensee')[['Total Gross Gaming Revenue', 'Patron_
↳ Winnings', 'Wagers']].mean().plot(kind='bar', figsize=(10, 6))
```

```
[ ]: <Axes: xlabel='Licensee'>
```



```
[ ]: corr_matrix = numerical.corr()
corr_matrix
```

```
[ ]:
           Wagers  Patron Winnings  \
Wagers          1.000000          0.999986
Patron Winnings  0.999986          1.000000
Cancelled Wagers  0.792654          0.791001
Online Casino Gaming Win/(Loss)  0.988162          0.987337
Promotional Coupons or Credits Wagered (3)  0.529160          0.530883
Promotional Deduction (4)          0.968887          0.968870
Total Gross Gaming Revenue          0.982202          0.981223
Payment (5)          0.982197          0.981217

           Cancelled Wagers  \
```


Wagers	0.792654
Patron Winnings	0.791001
Cancelled Wagers	1.000000
Online Casino Gaming Win/(Loss)	0.821653
Promotional Coupons or Credits Wagered (3)	0.286045
Promotional Deduction (4)	0.726616
Total Gross Gaming Revenue	0.832672
Payment (5)	0.832695

	Online Casino Gaming Win/(Loss) \
Wagers	0.988162
Patron Winnings	0.987337
Cancelled Wagers	0.821653
Online Casino Gaming Win/(Loss)	1.000000
Promotional Coupons or Credits Wagered (3)	0.474272
Promotional Deduction (4)	0.958824
Total Gross Gaming Revenue	0.998350
Payment (5)	0.998348

	Promotional Coupons or Credits
Wagered (3) \	
Wagers	
0.529160	
Patron Winnings	
0.530883	
Cancelled Wagers	
0.286045	
Online Casino Gaming Win/(Loss)	
0.474272	
Promotional Coupons or Credits Wagered (3)	
1.000000	
Promotional Deduction (4)	
0.616282	
Total Gross Gaming Revenue	
0.440832	
Payment (5)	
0.440810	

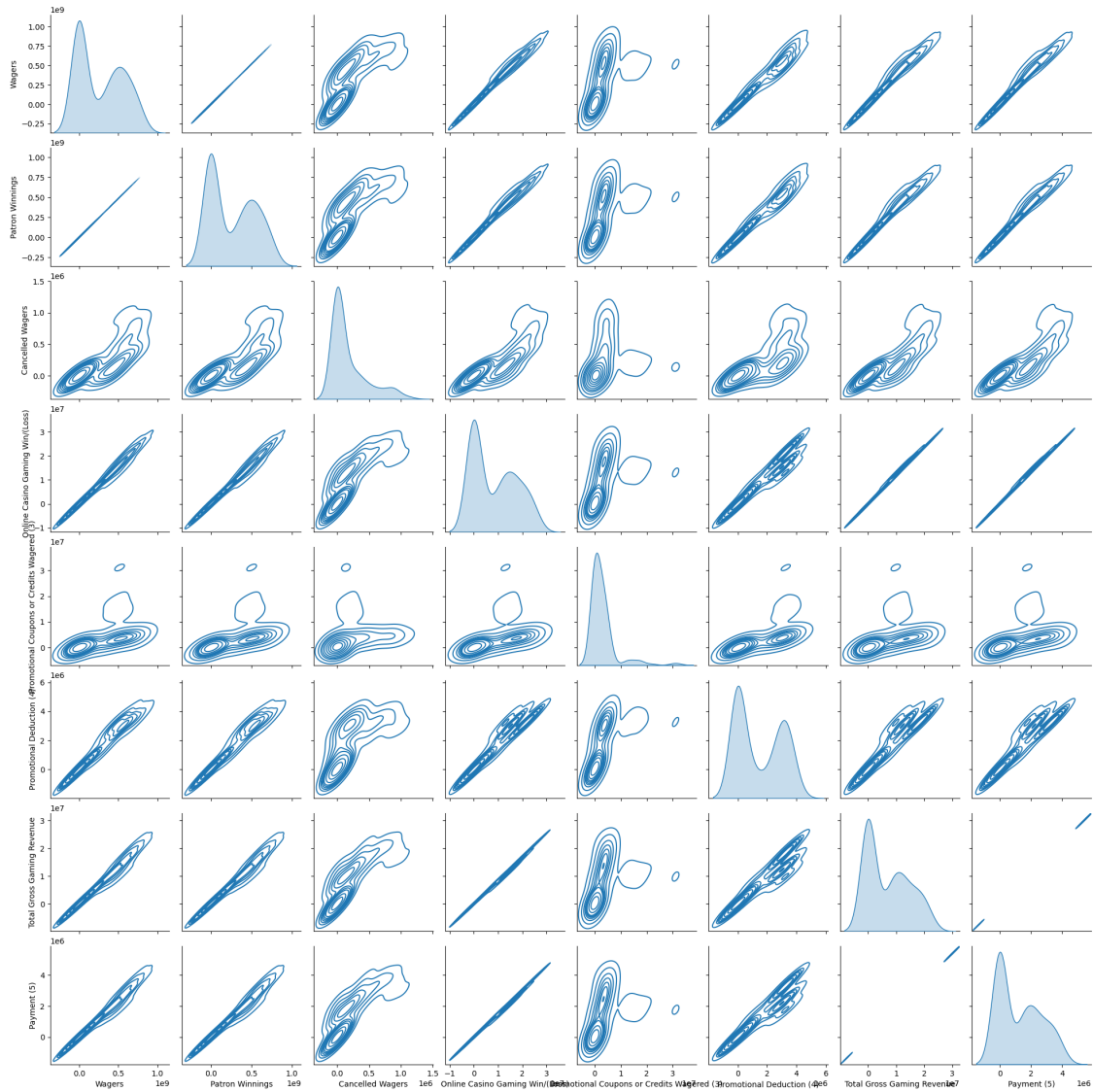
	Promotional Deduction (4) \
Wagers	0.968887
Patron Winnings	0.968870
Cancelled Wagers	0.726616
Online Casino Gaming Win/(Loss)	0.958824
Promotional Coupons or Credits Wagered (3)	0.616282
Promotional Deduction (4)	1.000000
Total Gross Gaming Revenue	0.940936
Payment (5)	0.940924

	Total Gross Gaming Revenue \
Wagers	0.982202
Patron Winnings	0.981223
Cancelled Wagers	0.832672
Online Casino Gaming Win/(Loss)	0.998350
Promotional Coupons or Credits Wagered (3)	0.440832
Promotional Deduction (4)	0.940936
Total Gross Gaming Revenue	1.000000
Payment (5)	0.999999

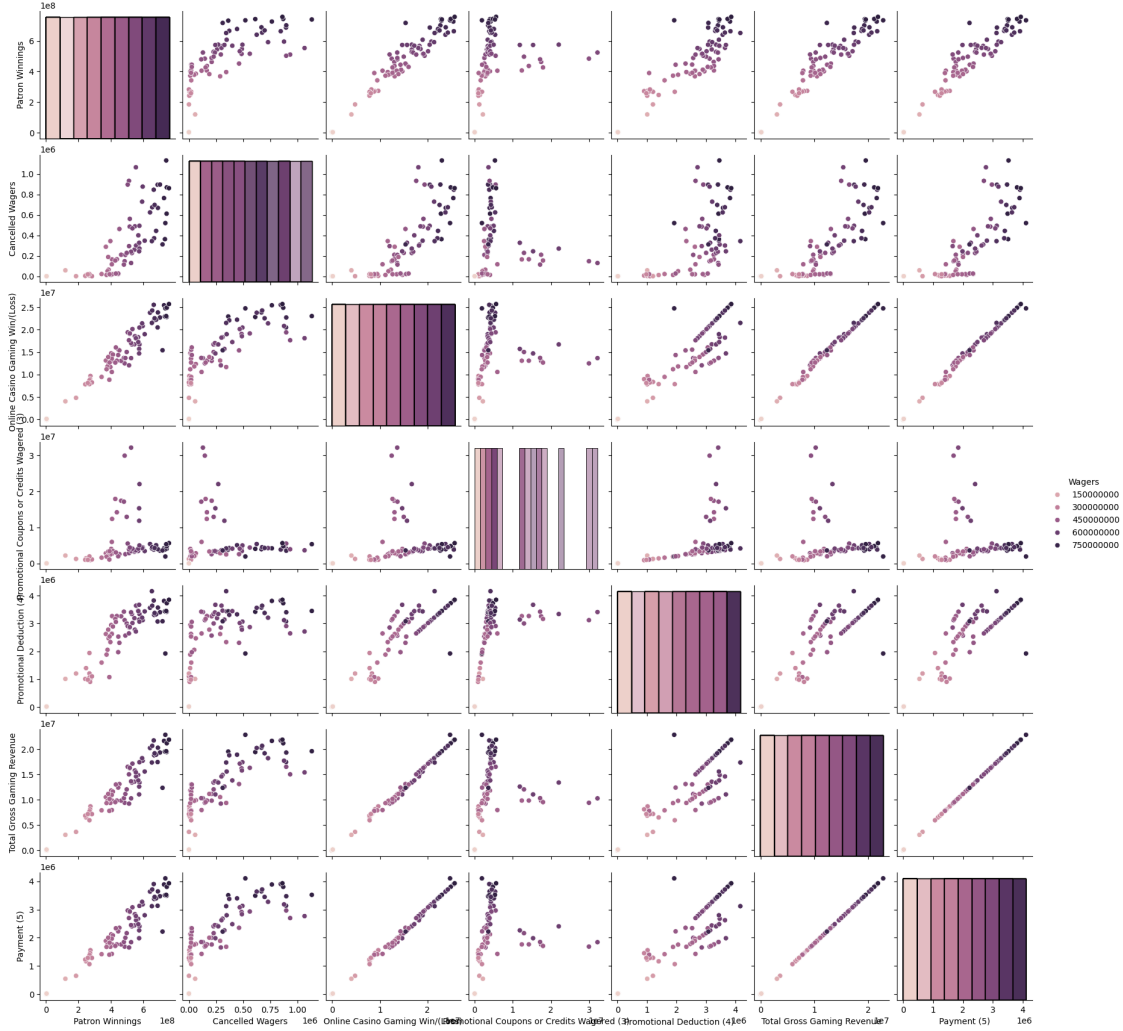
	Payment (5)
Wagers	0.982197
Patron Winnings	0.981217
Cancelled Wagers	0.832695
Online Casino Gaming Win/(Loss)	0.998348
Promotional Coupons or Credits Wagered (3)	0.440810
Promotional Deduction (4)	0.940924
Total Gross Gaming Revenue	0.999999
Payment (5)	1.000000

0.2 INSIGHTS

```
[ ]: sns.pairplot(numerical, kind='kde')
plt.show()
```



```
[ ]: sns.pairplot(numerical, hue ="Wagers", diag_kind='hist')
plt.show()
```



- There are two underperforming casinos according to boxplot. Mohegan Tribe and MPTN On-Reservation. MPI Master Wagering generates the highest value.
- As number of wagers increases, the gross revenue increases too. By monthly trends a pattern of increase and decrease to be observed. The wagers and the total gross gaming revenue increased through the years.
- Promotional credits and deductions increased the revenue. There is a 0.97 correlation with deductions and total revenue.
- As the wagers increases, the patron winnings also increase, and casino win/loss is also determined by the wagers.
- Promotional coupons and total gaming gross revenue has a correlation of 0.44, which would mean a negative impact on the total gross gaming revenue.
- Cancelled wagers has a negative impact on gross gaming revenue. As it increased the correlation to 0.83 from 0.98 which was the correlation with the normal total wagers.
- Key business insight would be: according to revenue patterns and correlations from the dataset given, the casinos should adjust the wagers accordingly to maximize the revenue. The patterns are observed through the months and it can be seen that a threefold sized

ladder pattern first increase and decrease with mode 3. So that casinos can choose to adjust the wagers.

- Patron winnings and payments go hand in hand, total gross gaming revenue and online casino wins and losses too.

```
[ ]: # linear regression
```

```
[71]: from sklearn.model_selection import train_test_split
      from sklearn.linear_model import LinearRegression
      from sklearn.metrics import mean_squared_error, r2_score

      X = df[['Wagers', 'Cancelled Wagers', 'Patron Winnings', 'Promotional Coupons_
      ↳or Credits Wagered (3)', 'Promotional Deduction (4)']]
      y = df['Total Gross Gaming Revenue']

      X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=2,
      ↳random_state=42)

      model = LinearRegression()
      model.fit(X_train, y_train)

      y_pred = model.predict(X_test)

      mse = mean_squared_error(y_test, y_pred)
      r2 = r2_score(y_test, y_pred)

      print(f"Mean Squared Error: {mse}")
      print(f"R-squared: {r2}")
```

Mean Squared Error: 0.49984715591719575

R-squared: 0.9999999999999462