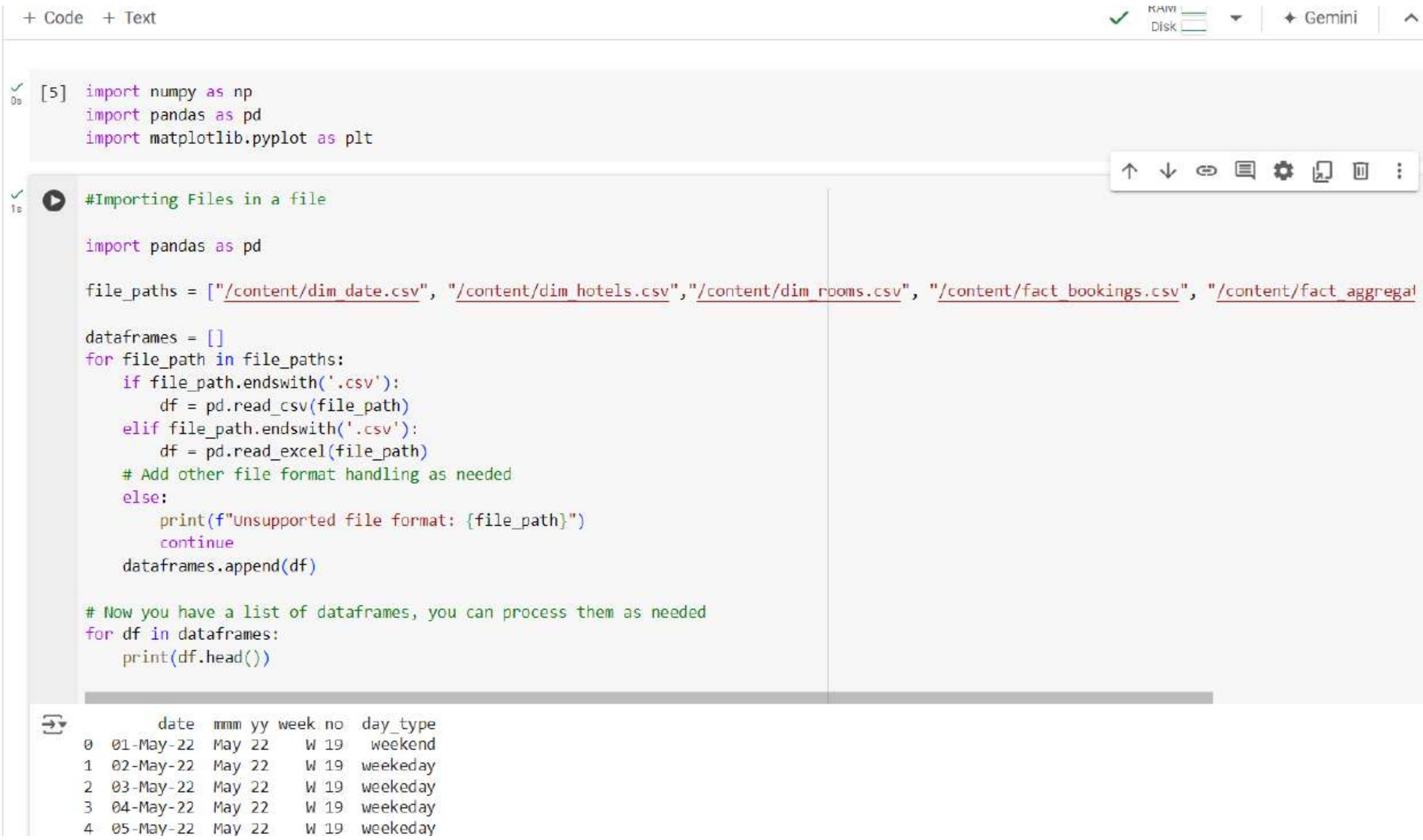
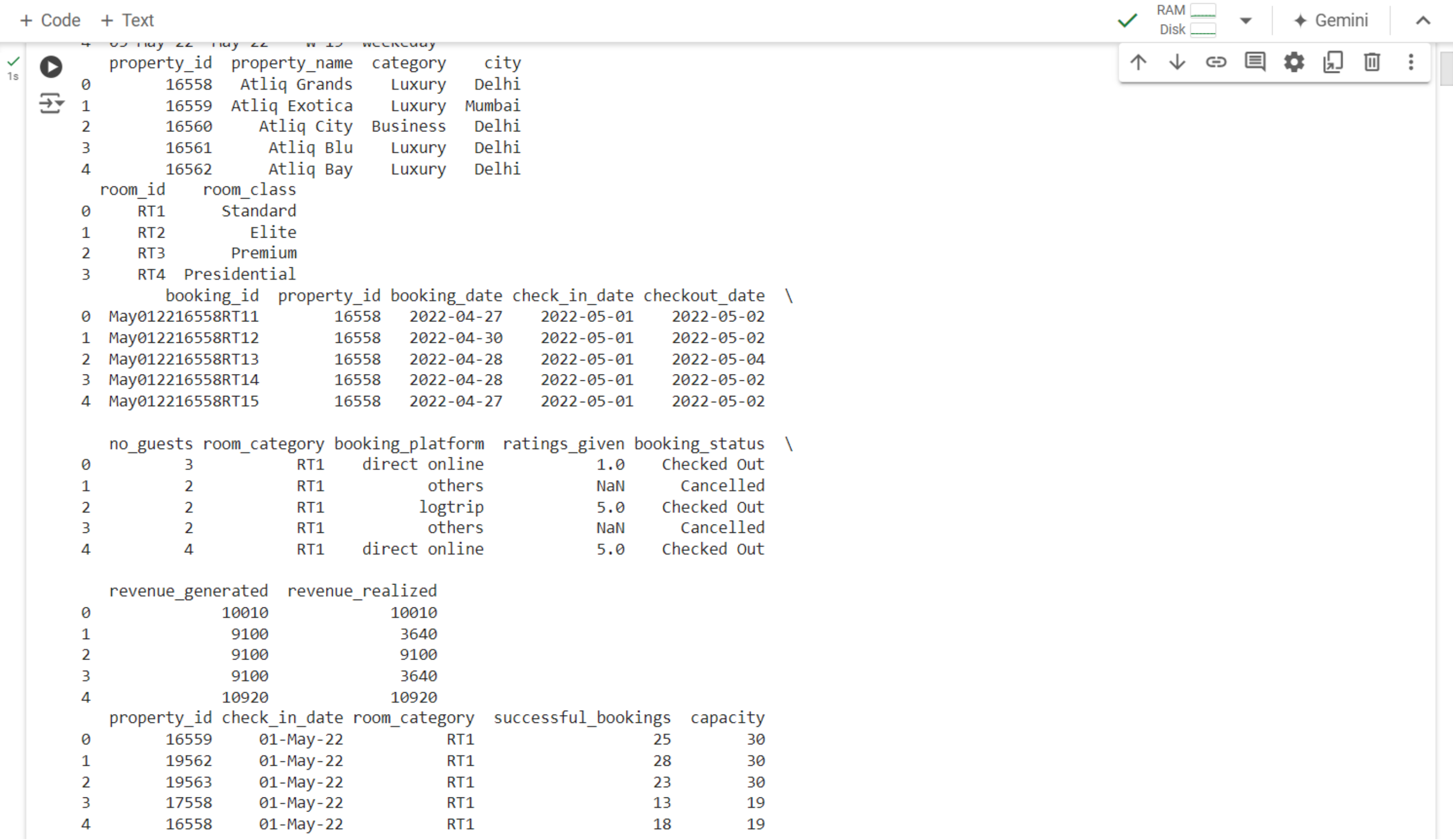


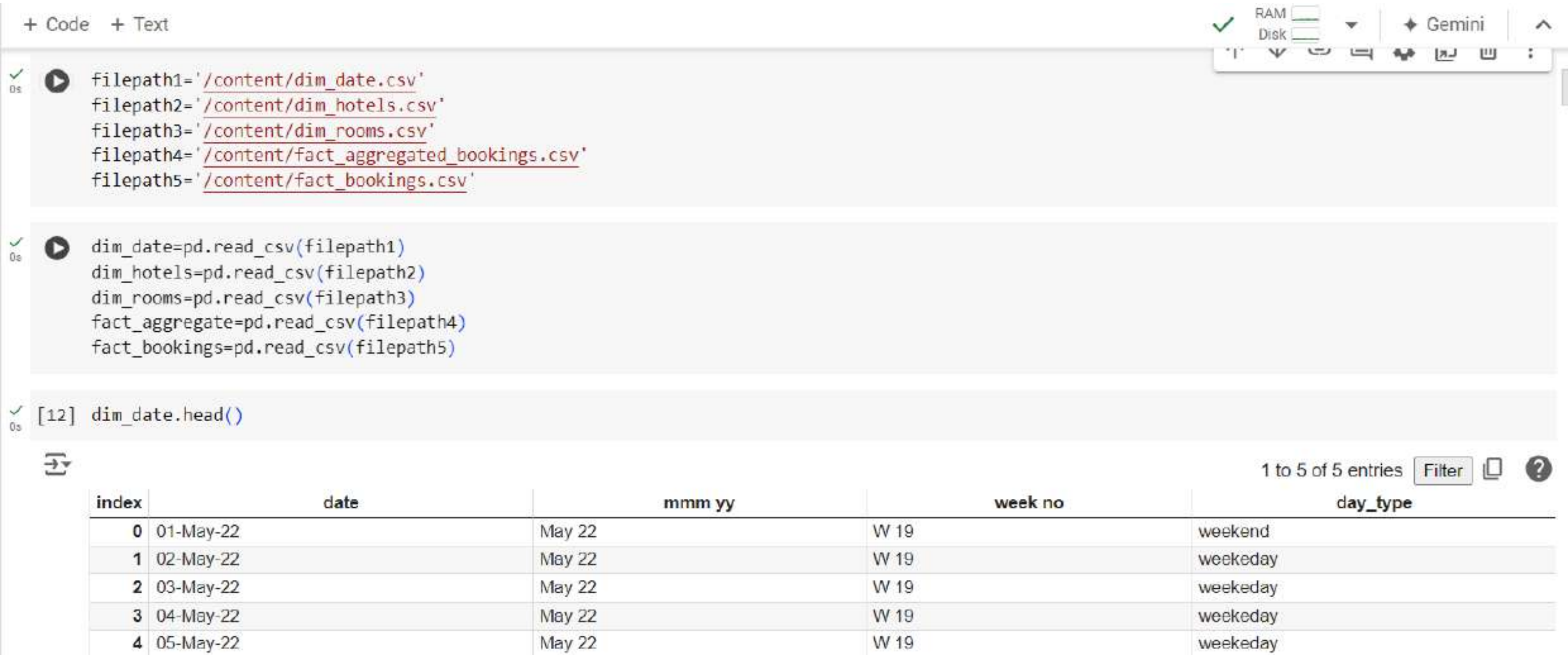
```
[5] import numpy as np
    import pandas as pd
    import matplotlib.pyplot as plt
```

```
#Importing Files in a file

import pandas as pd

file_paths = ["/content/dim_date.csv", "/content/dim_hotels.csv","/content/dim_rooms.csv", "/content/fact_bookings.csv", "/content/fact_aggregat

dataframes = []
for file_path in file_paths:
    if file_path.endswith('.csv'):
        df = pd.read_csv(file_path)
    elif file_path.endswith('.csv'):
        df = pd.read_excel(file_path)
    # Add other file format handling as needed
    else:
        print(f"Unsupported file format: {file_path}")
        continue
    dataframes.append(df)

# Now you have a list of dataframes, you can process them as needed
for df in dataframes:
    print(df.head())
```

```
        date  mmm yy week no  day_type
0  01-May-22  May 22    W 19   weekend
1  02-May-22  May 22    W 19  weekeday
2  03-May-22  May 22    W 19  weekeday
3  04-May-22  May 22    W 19  weekeday
4  05-May-22  May 22    W 19  weekeday
```



```
4  05-May-22  May 22    W 19  weekday
   property_id  property_name  category    city
0        16558   Atliq Grands    Luxury   Delhi
1        16559  Atliq Exotica    Luxury  Mumbai
2        16560     Atliq City  Business   Delhi
3        16561      Atliq Blu    Luxury   Delhi
4        16562      Atliq Bay    Luxury   Delhi
  room_id    room_class
0     RT1      Standard
1     RT2         Elite
2     RT3       Premium
3     RT4  Presidential
         booking_id  property_id booking_date check_in_date checkout_date  \
0  May012216558RT11        16558   2022-04-27    2022-05-01    2022-05-02
1  May012216558RT12        16558   2022-04-30    2022-05-01    2022-05-02
2  May012216558RT13        16558   2022-04-28    2022-05-01    2022-05-04
3  May012216558RT14        16558   2022-04-28    2022-05-01    2022-05-02
4  May012216558RT15        16558   2022-04-27    2022-05-01    2022-05-02

   no_guests room_category booking_platform  ratings_given booking_status  \
0          3           RT1    direct online            1.0    Checked Out
1          2           RT1           others            NaN      Cancelled
2          2           RT1          logtrip            5.0    Checked Out
3          2           RT1           others            NaN      Cancelled
4          4           RT1    direct online            5.0    Checked Out

   revenue_generated  revenue_realized
0              10010             10010
1               9100              3640
2               9100              9100
3               9100              3640
4              10920             10920
   property_id check_in_date room_category  successful_bookings  capacity
0        16559     01-May-22           RT1                   25        30
1        19562     01-May-22           RT1                   28        30
2        19563     01-May-22           RT1                   23        30
3        17558     01-May-22           RT1                   13        19
4        16558     01-May-22           RT1                   18        19
```

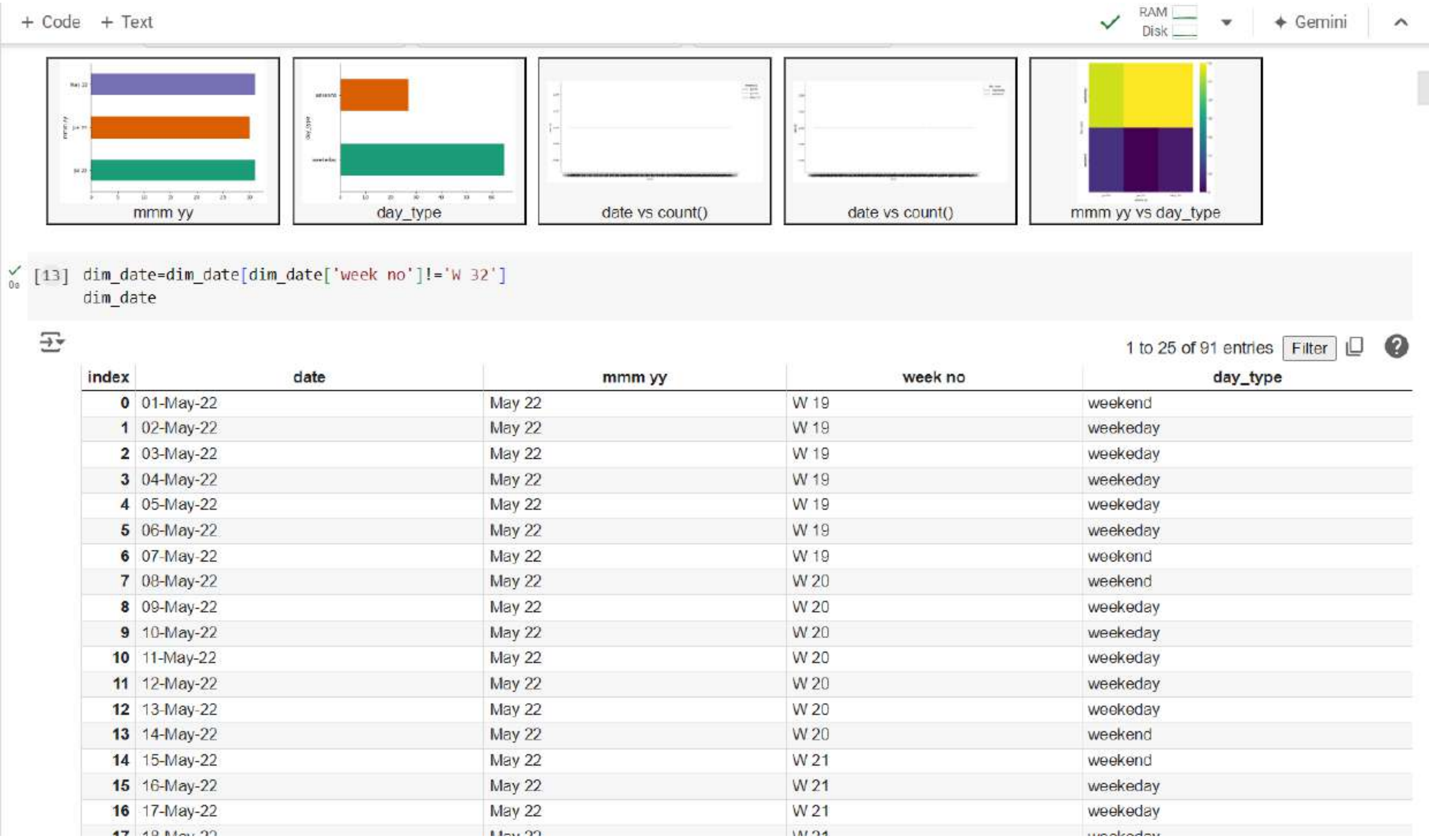


```
filepath1='/content/dim_date.csv'
filepath2='/content/dim_hotels.csv'
filepath3='/content/dim_rooms.csv'
filepath4='/content/fact_aggregated_bookings.csv'
filepath5='/content/fact_bookings.csv'
```

```
dim_date=pd.read_csv(filepath1)
dim_hotels=pd.read_csv(filepath2)
dim_rooms=pd.read_csv(filepath3)
fact_aggregate=pd.read_csv(filepath4)
fact_bookings=pd.read_csv(filepath5)
```

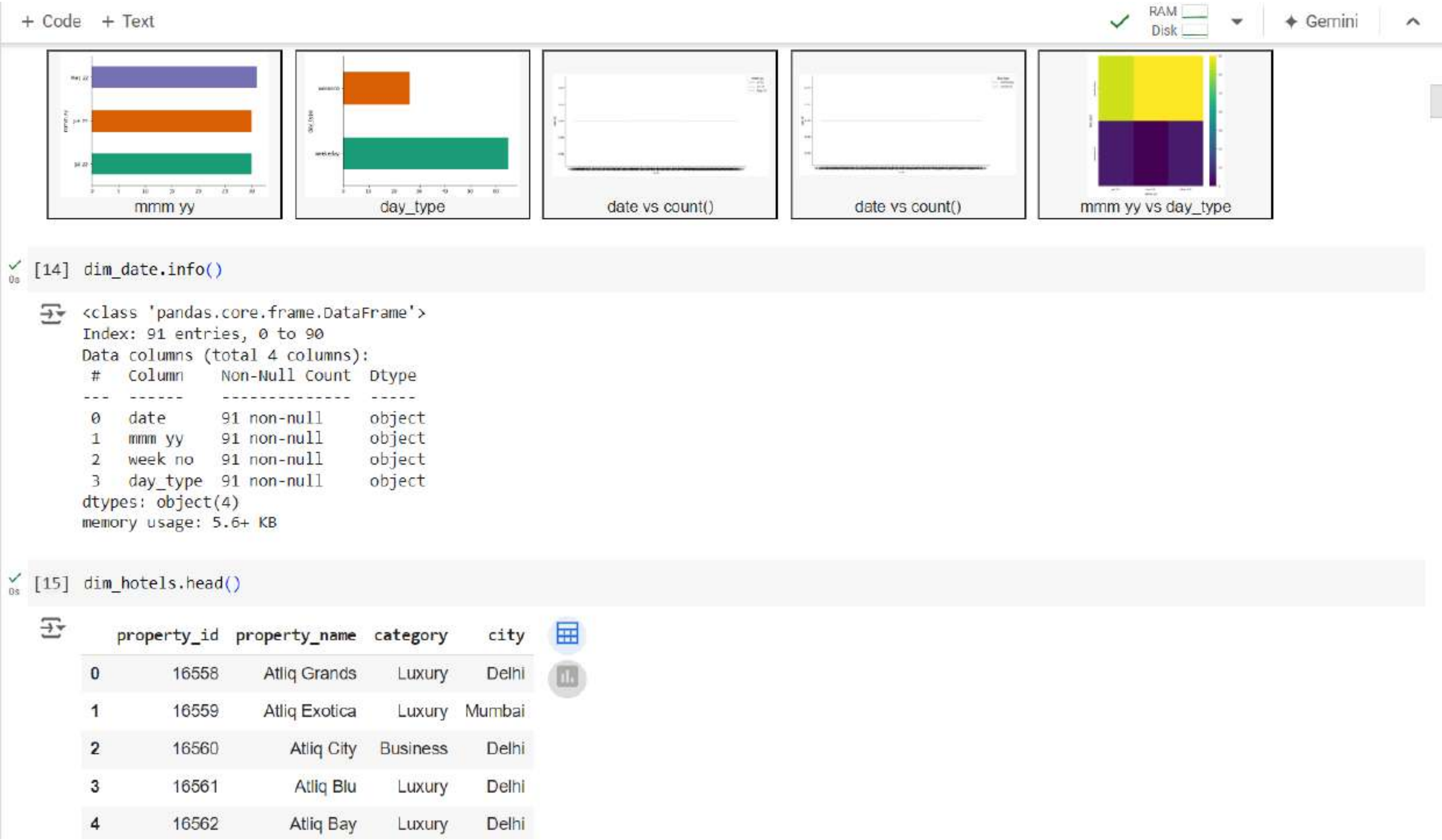
```
[12] dim_date.head()
```

index	date	mmm yy	week no	day_type
0	01-May-22	May 22	W 19	weekend
1	02-May-22	May 22	W 19	weekeday
2	03-May-22	May 22	W 19	weekeday
3	04-May-22	May 22	W 19	weekeday
4	05-May-22	May 22	W 19	weekeday



```
[13] dim_date=dim_date[dim_date['week no']!='W 32']
     dim_date
```

index	date	mmm yy	week no	day_type
0	01-May-22	May 22	W 19	weekend
1	02-May-22	May 22	W 19	weekeday
2	03-May-22	May 22	W 19	weekeday
3	04-May-22	May 22	W 19	weekeday
4	05-May-22	May 22	W 19	weekeday
5	06-May-22	May 22	W 19	weekeday
6	07-May-22	May 22	W 19	weekend
7	08-May-22	May 22	W 20	weekend
8	09-May-22	May 22	W 20	weekeday
9	10-May-22	May 22	W 20	weekeday
10	11-May-22	May 22	W 20	weekeday
11	12-May-22	May 22	W 20	weekeday
12	13-May-22	May 22	W 20	weekeday
13	14-May-22	May 22	W 20	weekend
14	15-May-22	May 22	W 21	weekend
15	16-May-22	May 22	W 21	weekeday
16	17-May-22	May 22	W 21	weekeday
17	18-May-22	May 22	W 21	weekeday

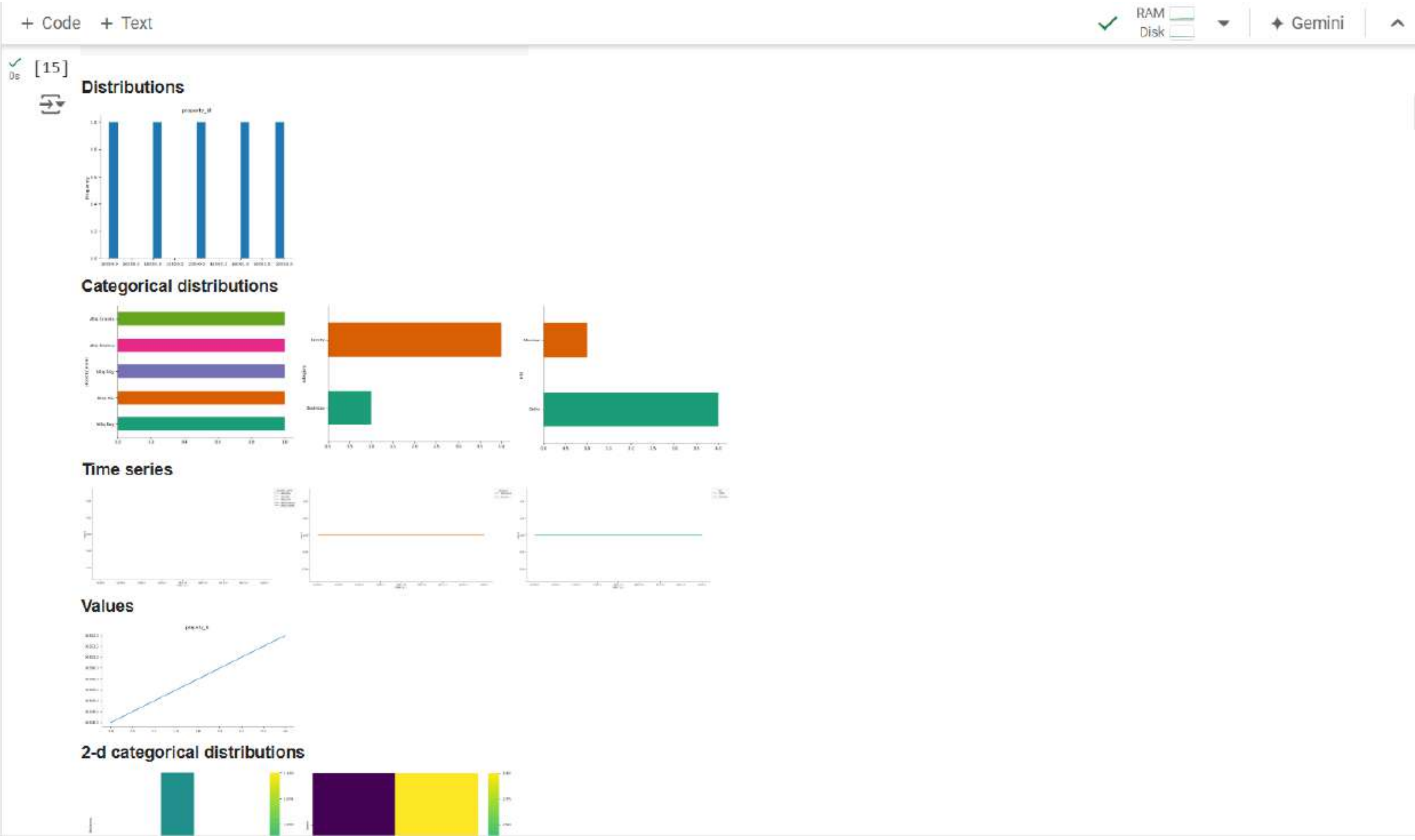


```
[14] dim_date.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 91 entries, 0 to 90
Data columns (total 4 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   date      91 non-null     object
 1   mmm yy    91 non-null     object
 2   week no   91 non-null     object
 3   day_type  91 non-null     object
dtypes: object(4)
memory usage: 5.6+ KB
```

```
[15] dim_hotels.head()
```

	property_id	property_name	category	city
0	16558	Atliq Grands	Luxury	Delhi
1	16559	Atliq Exotica	Luxury	Mumbai
2	16560	Atliq City	Business	Delhi
3	16561	Atliq Blu	Luxury	Delhi
4	16562	Atliq Bay	Luxury	Delhi



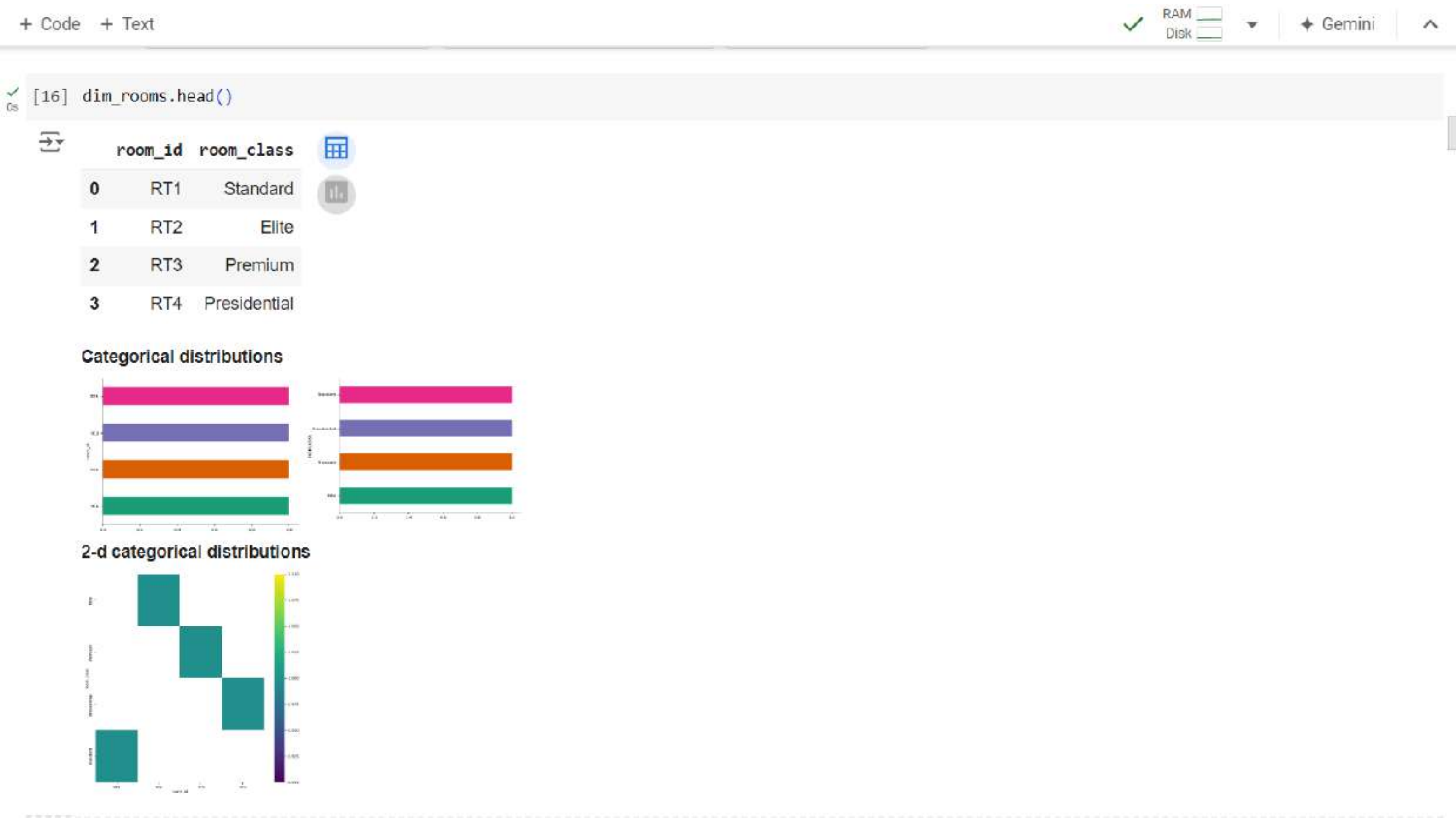
Distributions

Categorical distributions

Time series

Values

2-d categorical distributions

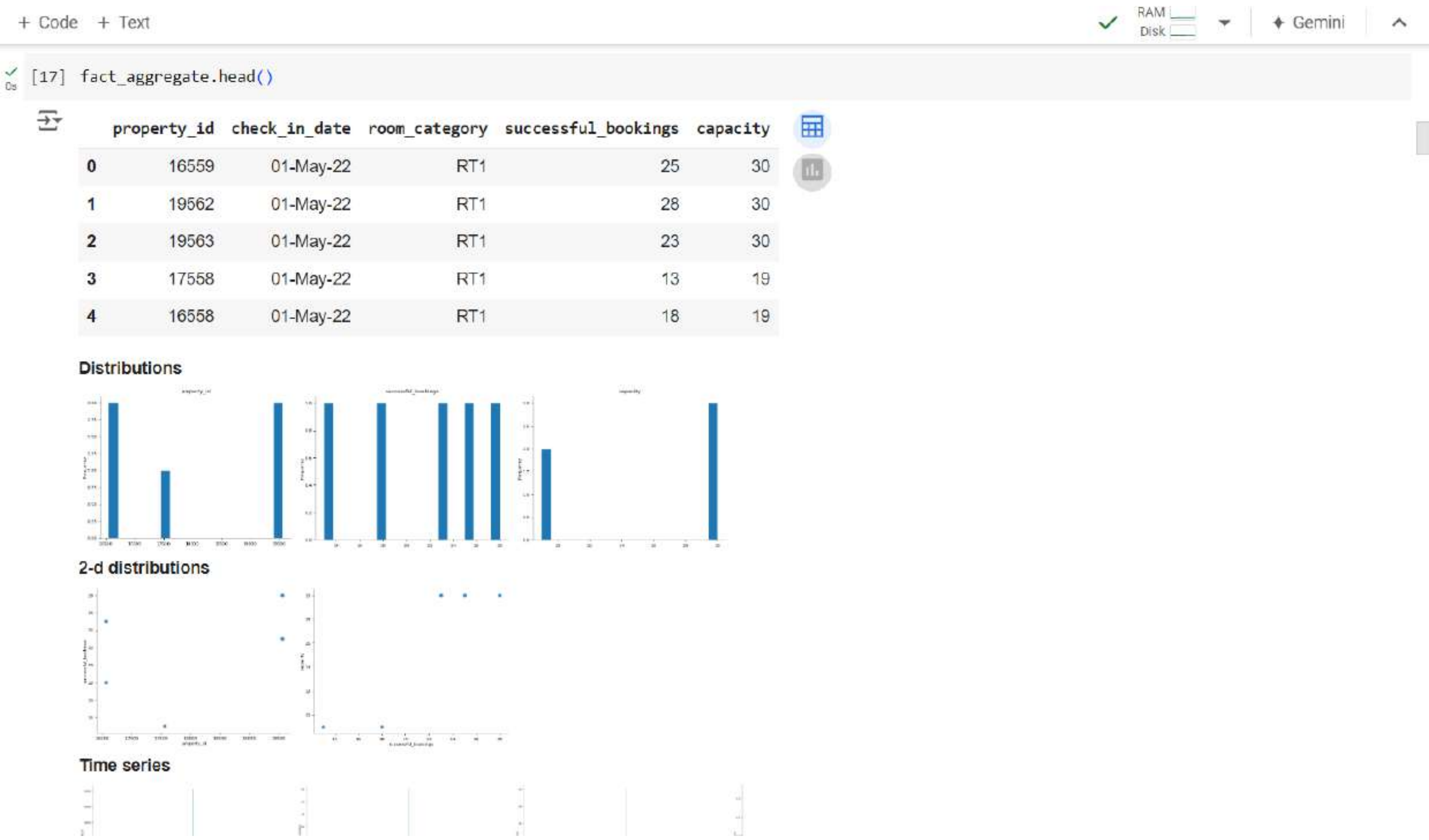


```
[16] dim_rooms.head()
```

	room_id	room_class
0	RT1	Standard
1	RT2	Elite
2	RT3	Premium
3	RT4	Presidential

Categorical distributions

2-d categorical distributions



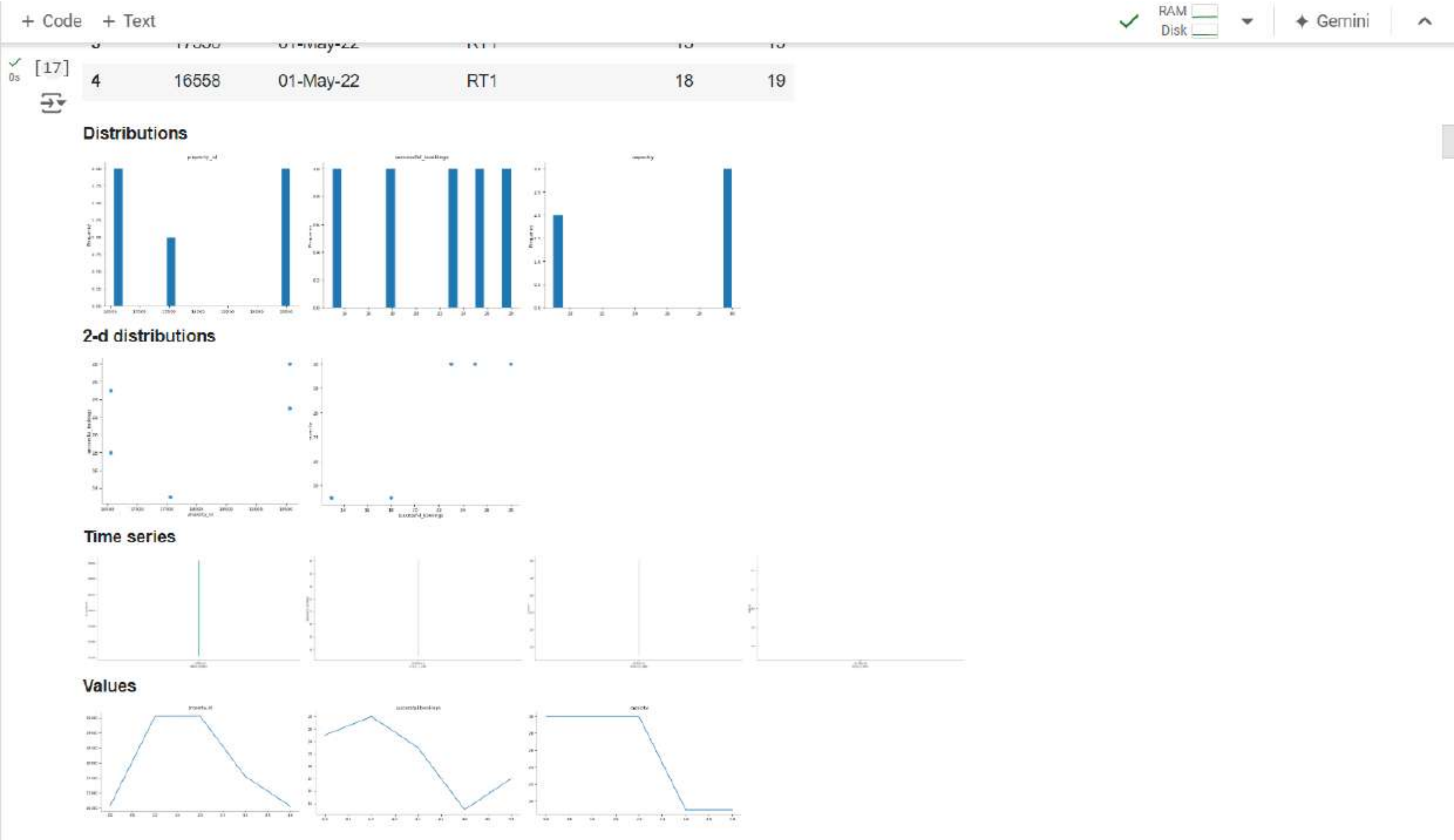
```
[17] fact_aggregate.head()
```

	property_id	check_in_date	room_category	successful_bookings	capacity
0	16559	01-May-22	RT1	25	30
1	19562	01-May-22	RT1	28	30
2	19563	01-May-22	RT1	23	30
3	17558	01-May-22	RT1	13	19
4	16558	01-May-22	RT1	18	19

Distributions

2-d distributions

Time series



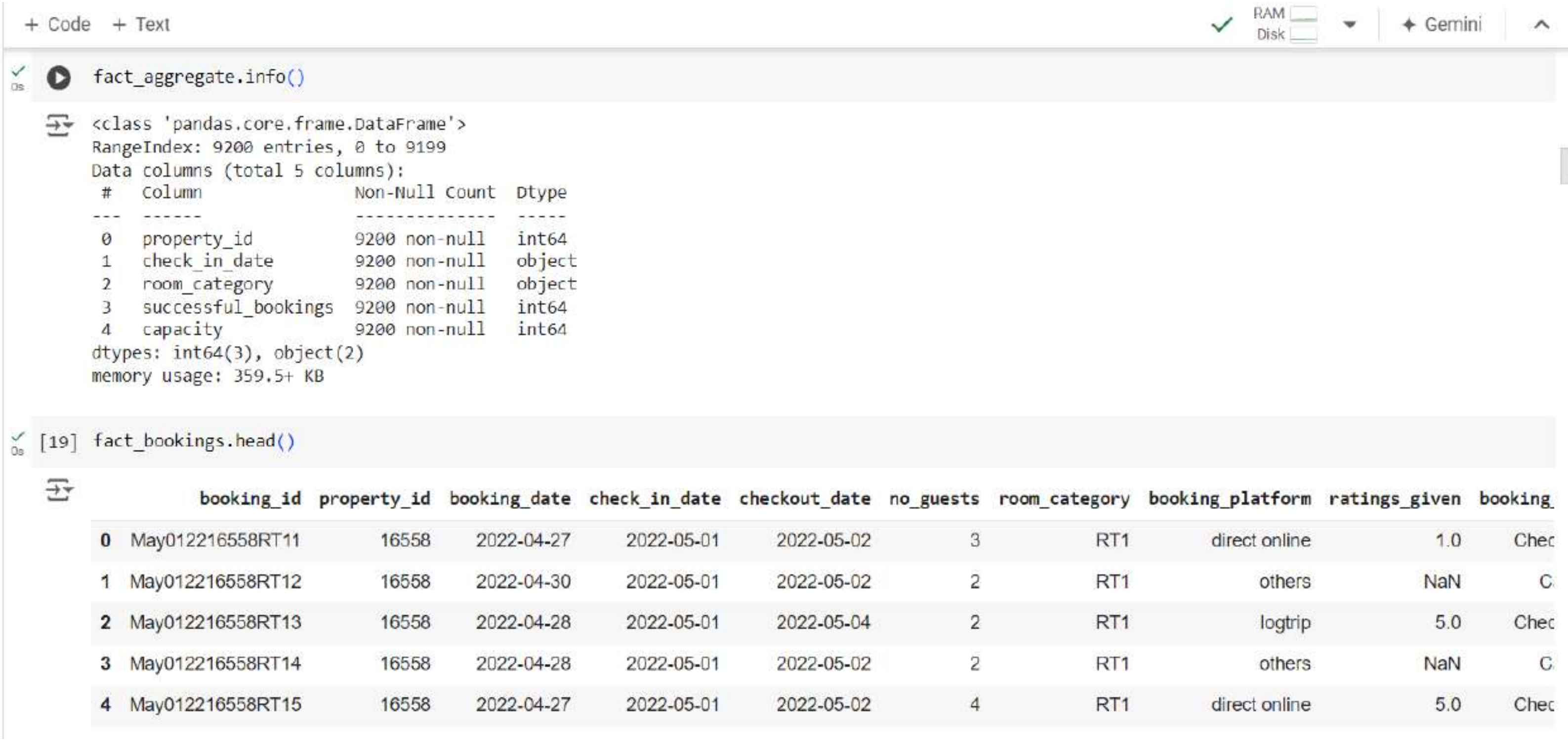
4	16558	01-May-22	RT1	18	19

Distributions

2-d distributions

Time series

Values

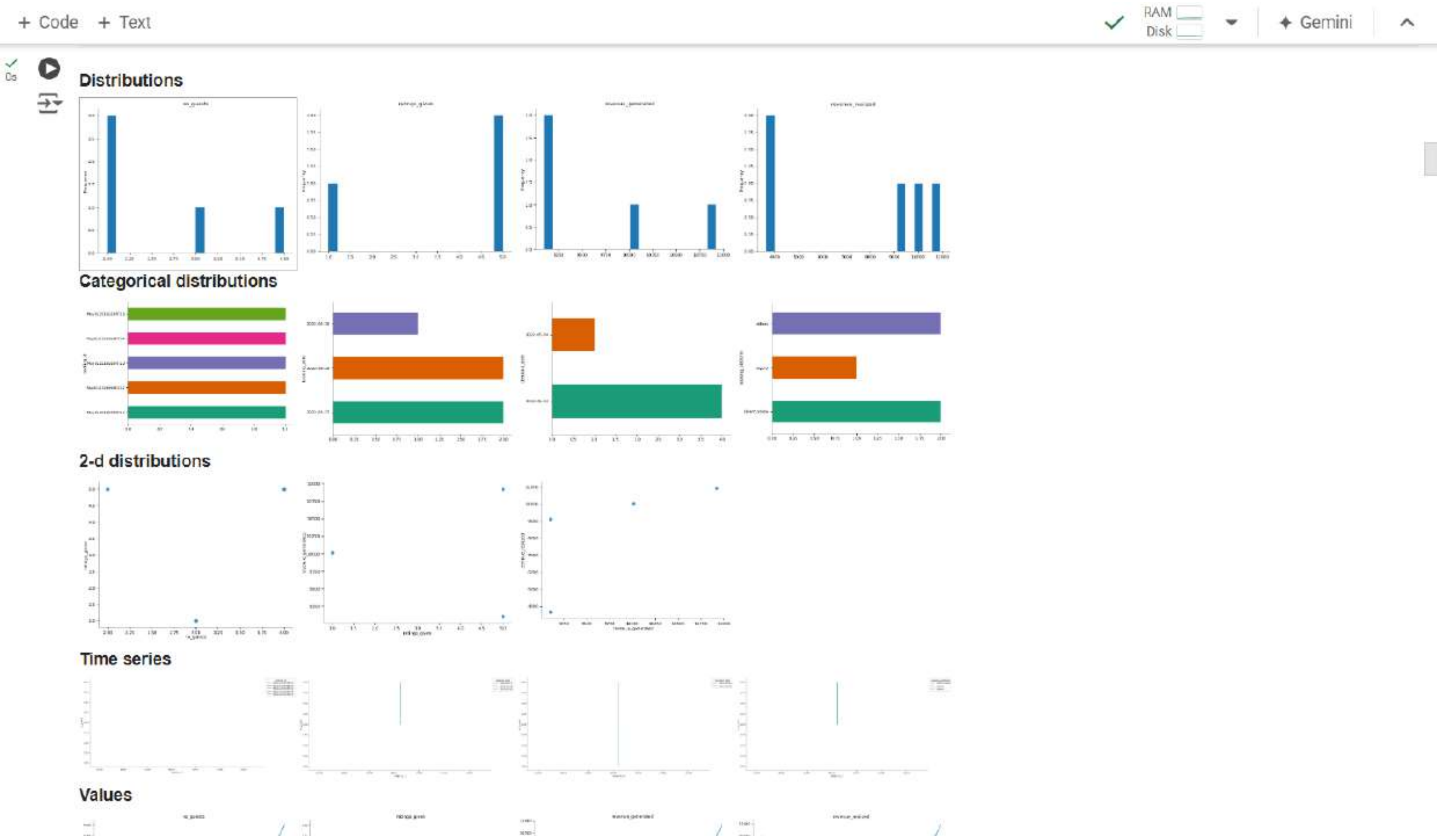


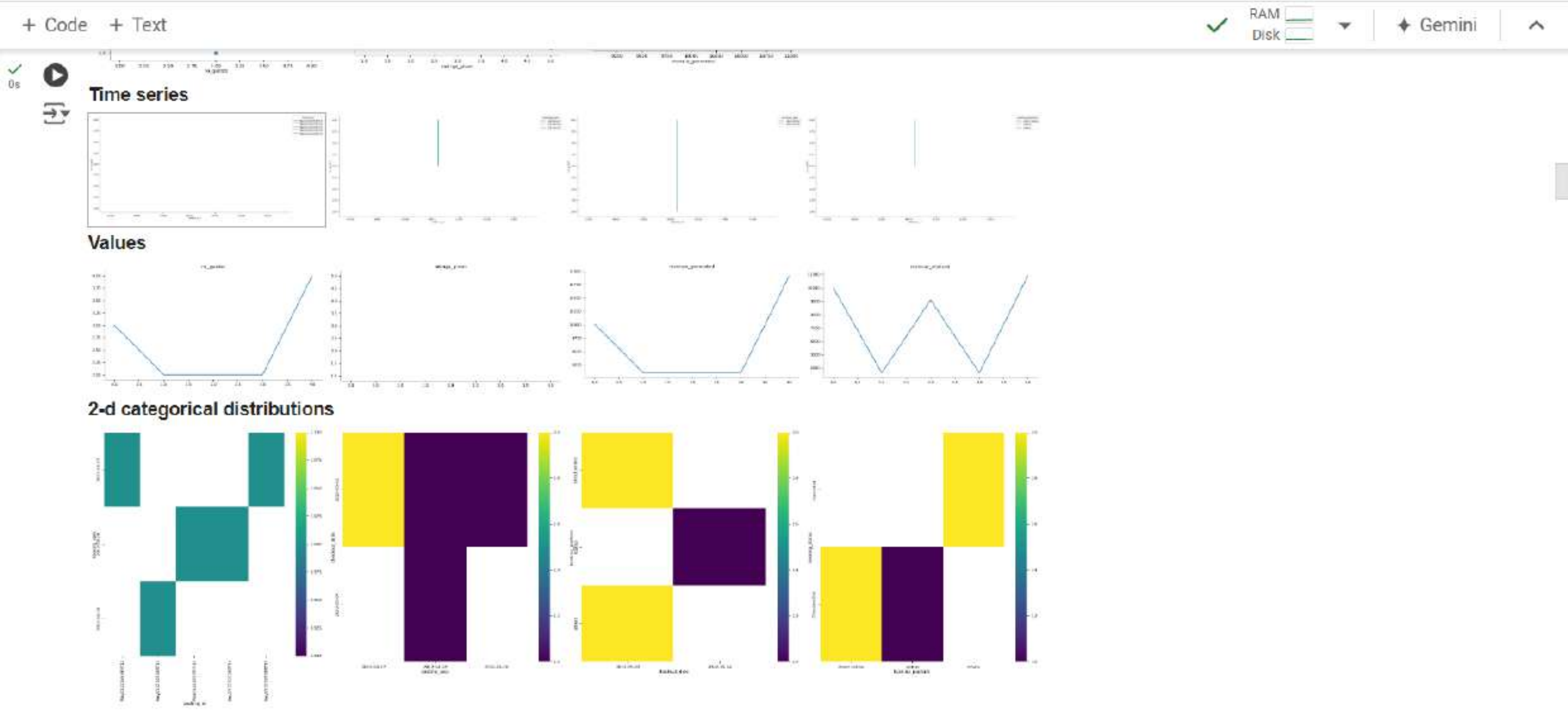
```
fact_aggregate.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9200 entries, 0 to 9199
Data columns (total 5 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   property_id          9200 non-null   int64
 1   check_in_date        9200 non-null   object
 2   room_category        9200 non-null   object
 3   successful_bookings  9200 non-null   int64
 4   capacity             9200 non-null   int64
dtypes: int64(3), object(2)
memory usage: 359.5+ KB
```

```
[19] fact_bookings.head()
```

	booking_id	property_id	booking_date	check_in_date	checkout_date	no_guests	room_category	booking_platform	ratings_given	booking_
0	May012216558RT11	16558	2022-04-27	2022-05-01	2022-05-02	3	RT1	direct online	1.0	Chec
1	May012216558RT12	16558	2022-04-30	2022-05-01	2022-05-02	2	RT1	others	NaN	C
2	May012216558RT13	16558	2022-04-28	2022-05-01	2022-05-04	2	RT1	logtrip	5.0	Chec
3	May012216558RT14	16558	2022-04-28	2022-05-01	2022-05-02	2	RT1	others	NaN	C
4	May012216558RT15	16558	2022-04-27	2022-05-01	2022-05-02	4	RT1	direct online	5.0	Chec

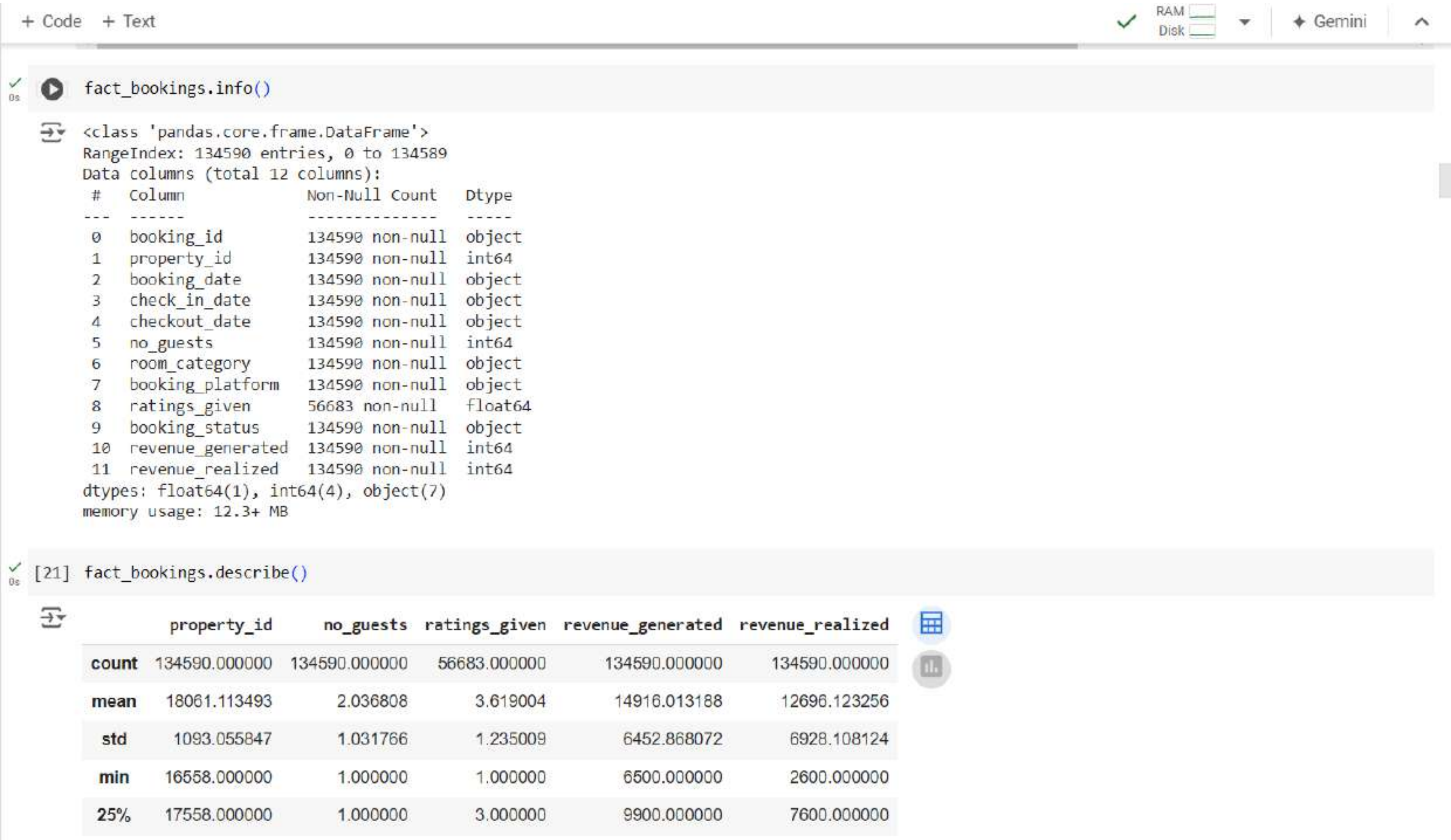

+ Code
+ Text
RAM
Disk
Gemini
Distributions
Categorical distributions
2-d distributions
Time series
Values



Time series

Values

2-d categorical distributions

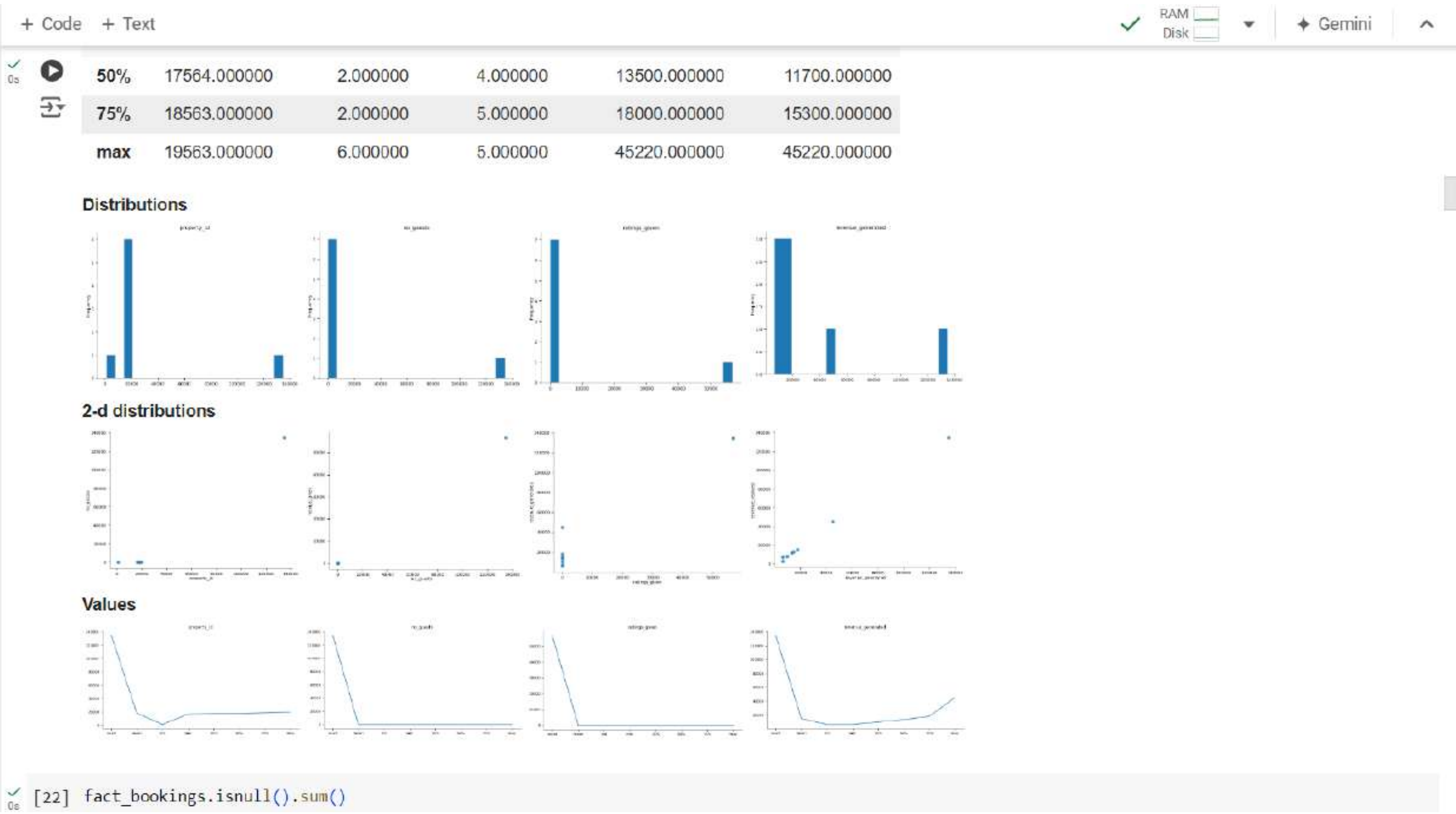


```
fact_bookings.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 134590 entries, 0 to 134589
Data columns (total 12 columns):
 #   Column             Non-Null Count   Dtype  
---  ------             --------------   -----  
 0   booking_id         134590 non-null  object 
 1   property_id        134590 non-null  int64  
 2   booking_date       134590 non-null  object 
 3   check_in_date      134590 non-null  object 
 4   checkout_date      134590 non-null  object 
 5   no_guests          134590 non-null  int64  
 6   room_category      134590 non-null  object 
 7   booking_platform   134590 non-null  object 
 8   ratings_given      56683 non-null   float64
 9   booking_status     134590 non-null  object 
 10  revenue_generated  134590 non-null  int64  
 11  revenue_realized   134590 non-null  int64  
dtypes: float64(1), int64(4), object(7)
memory usage: 12.3+ MB
```

```
[21] fact_bookings.describe()
```

	property_id	no_guests	ratings_given	revenue_generated	revenue_realized
count	134590.000000	134590.000000	56683.000000	134590.000000	134590.000000
mean	18061.113493	2.036808	3.619004	14916.013188	12696.123256
std	1093.055847	1.031766	1.235009	6452.868072	6928.108124
min	16558.000000	1.000000	1.000000	6500.000000	2600.000000
25%	17558.000000	1.000000	3.000000	9900.000000	7600.000000



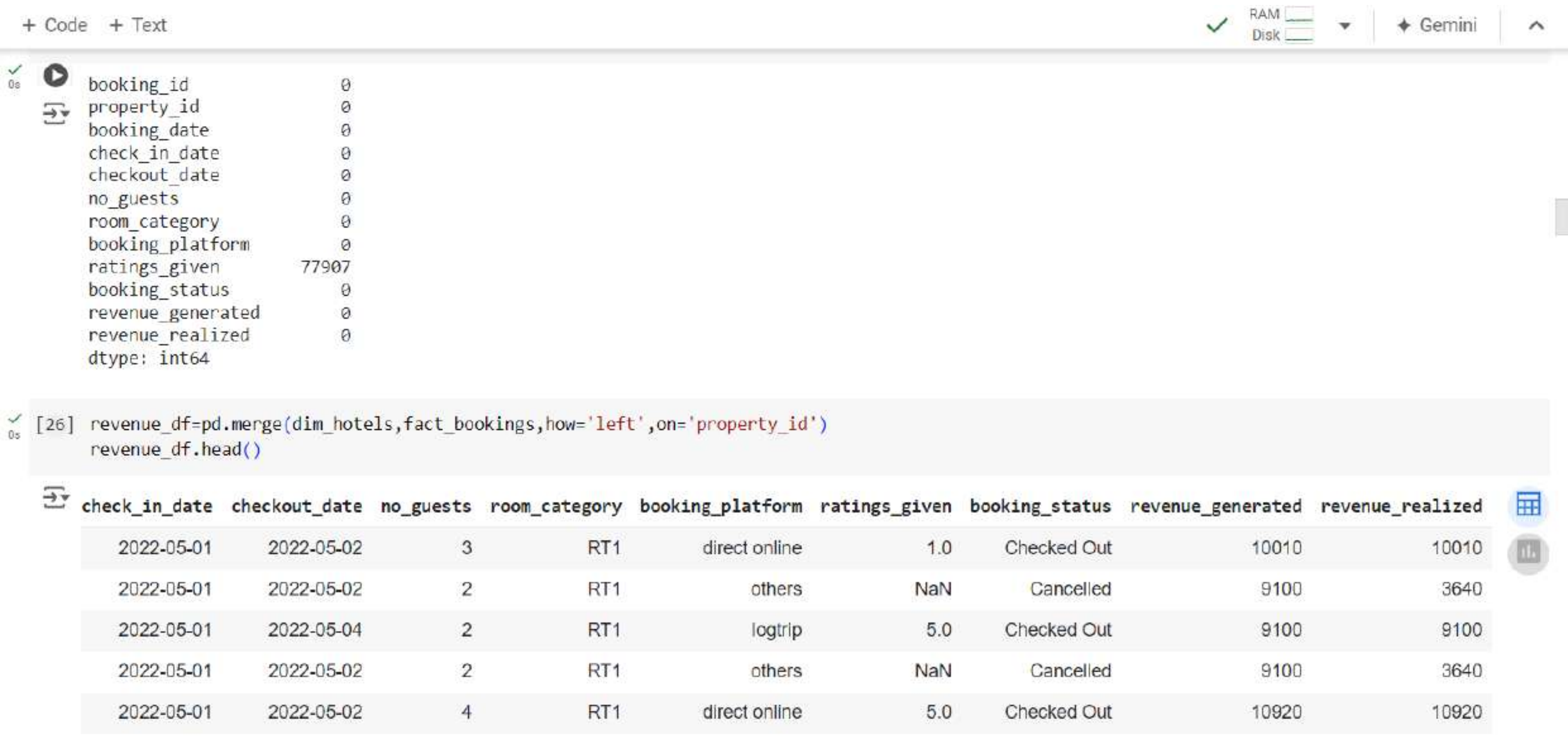
```
50%   17564.000000   2.000000   4.000000   13500.000000   11700.000000
75%   18563.000000   2.000000   5.000000   18000.000000   15300.000000
max   19563.000000   6.000000   5.000000   45220.000000   45220.000000
```

Distributions

2-d distributions

Values

```
[22] fact_bookings.isnull().sum()
```

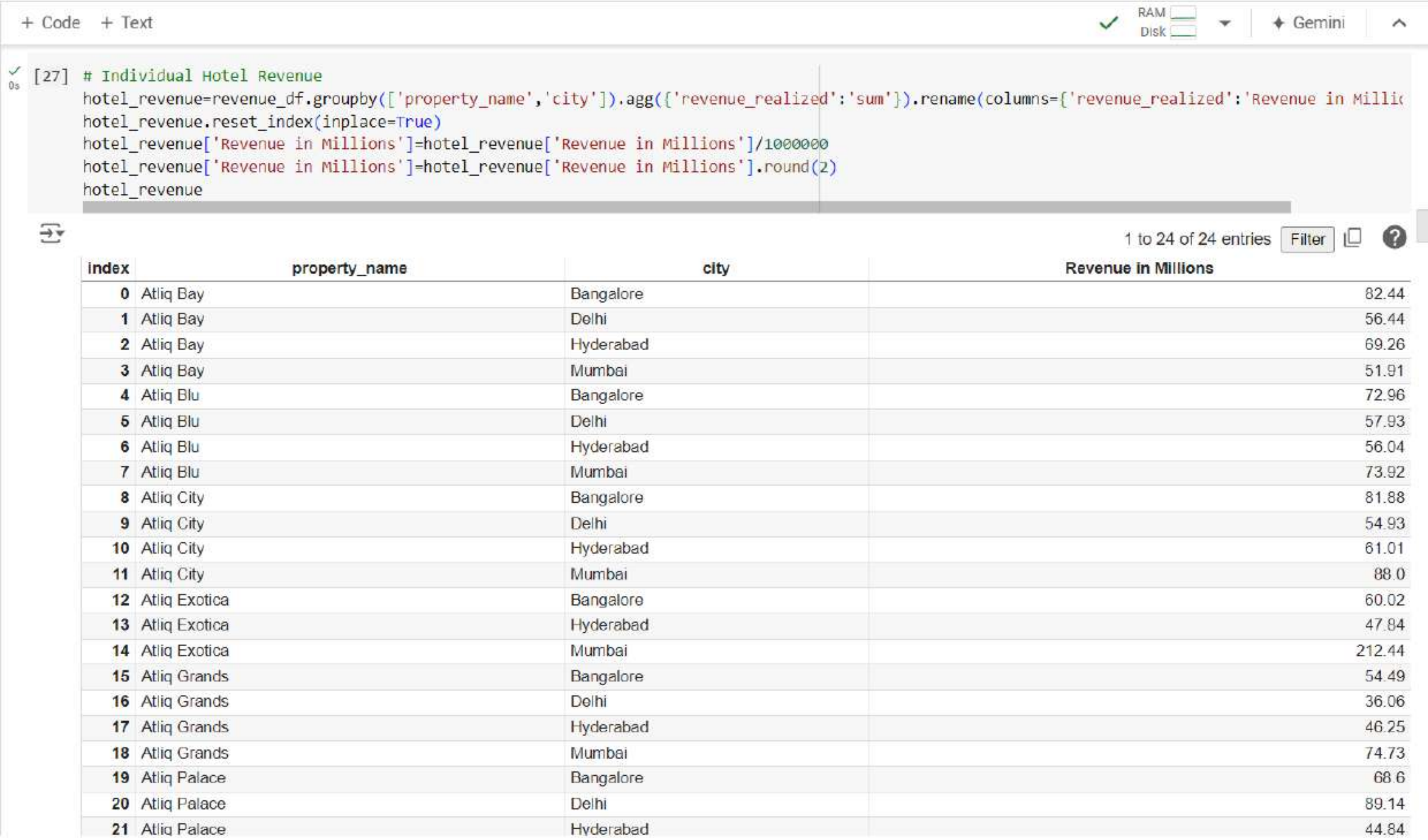


```
booking_id               0
property_id              0
booking_date             0
check_in_date            0
checkout_date            0
no_guests                0
room_category            0
booking_platform         0
ratings_given        77907
booking_status           0
revenue_generated        0
revenue_realized         0
dtype: int64
```

```
[26] revenue_df=pd.merge(dim_hotels,fact_bookings,how='left',on='property_id')
     revenue_df.head()
```

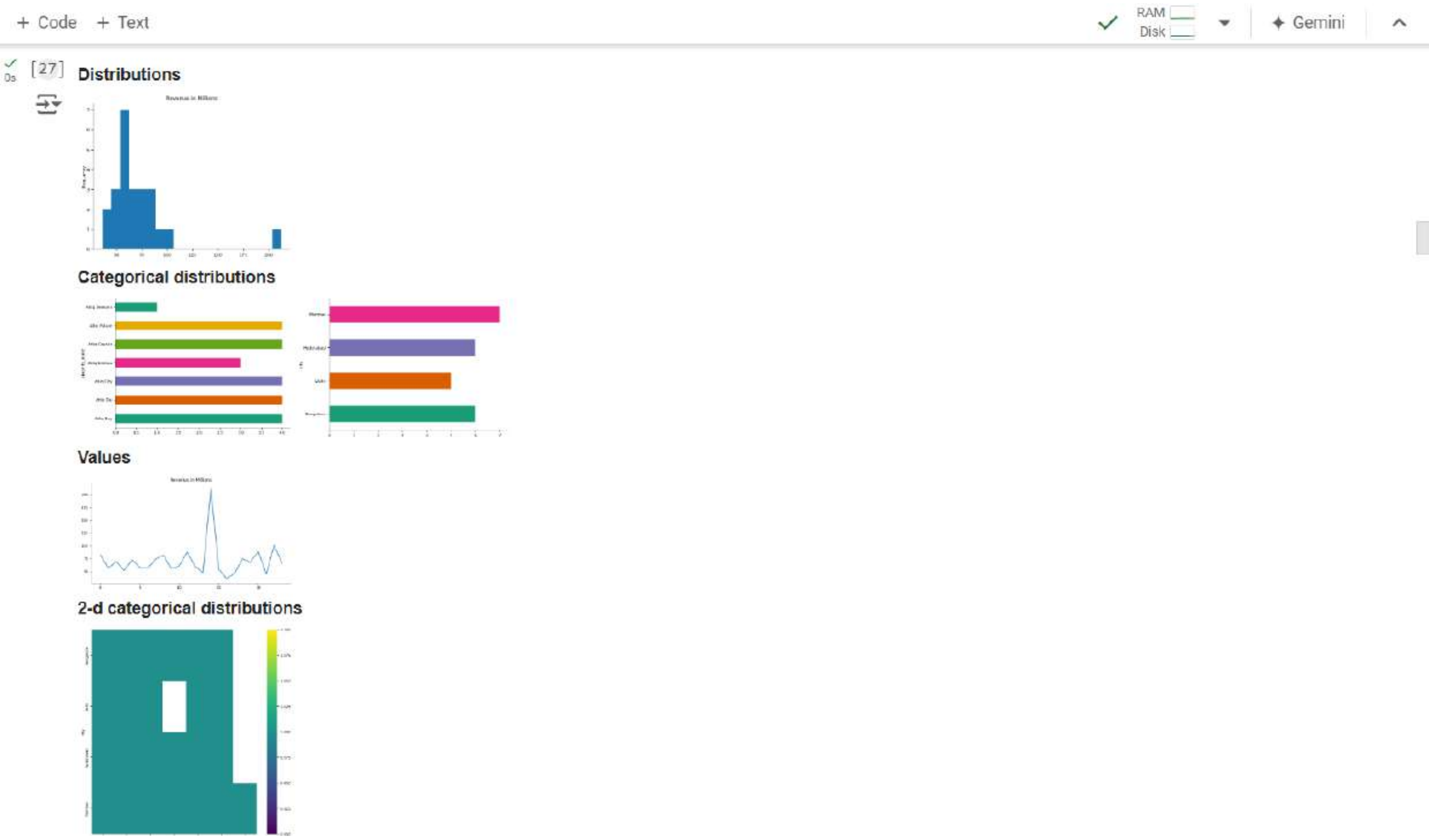
check_in_date	checkout_date	no_guests	room_category	booking_platform	ratings_given	booking_status	revenue_generated	revenue_realized
2022-05-01	2022-05-02	3	RT1	direct online	1.0	Checked Out	10010	10010
2022-05-01	2022-05-02	2	RT1	others	NaN	Cancelled	9100	3640
2022-05-01	2022-05-04	2	RT1	logtrip	5.0	Checked Out	9100	9100
2022-05-01	2022-05-02	2	RT1	others	NaN	Cancelled	9100	3640
2022-05-01	2022-05-02	4	RT1	direct online	5.0	Checked Out	10920	10920

revenue_realized



```
# Individual Hotel Revenue
hotel_revenue=revenue_df.groupby(['property_name','city']).agg({'revenue_realized':'sum'}).rename(columns={'revenue_realized':'Revenue in Millic
hotel_revenue.reset_index(inplace=True)
hotel_revenue['Revenue in Millions']=hotel_revenue['Revenue in Millions']/1000000
hotel_revenue['Revenue in Millions']=hotel_revenue['Revenue in Millions'].round(2)
hotel_revenue
```

index	property_name	city	Revenue in Millions
0	Atliq Bay	Bangalore	82.44
1	Atliq Bay	Delhi	56.44
2	Atliq Bay	Hyderabad	69.26
3	Atliq Bay	Mumbai	51.91
4	Atliq Blu	Bangalore	72.96
5	Atliq Blu	Delhi	57.93
6	Atliq Blu	Hyderabad	56.04
7	Atliq Blu	Mumbai	73.92
8	Atliq City	Bangalore	81.88
9	Atliq City	Delhi	54.93
10	Atliq City	Hyderabad	61.01
11	Atliq City	Mumbai	88.0
12	Atliq Exotica	Bangalore	60.02
13	Atliq Exotica	Hyderabad	47.84
14	Atliq Exotica	Mumbai	212.44
15	Atliq Grands	Bangalore	54.49
16	Atliq Grands	Delhi	36.06
17	Atliq Grands	Hyderabad	46.25
18	Atliq Grands	Mumbai	74.73
19	Atliq Palace	Bangalore	68.6
20	Atliq Palace	Delhi	89.14
21	Atliq Palace	Hyderabad	44.84

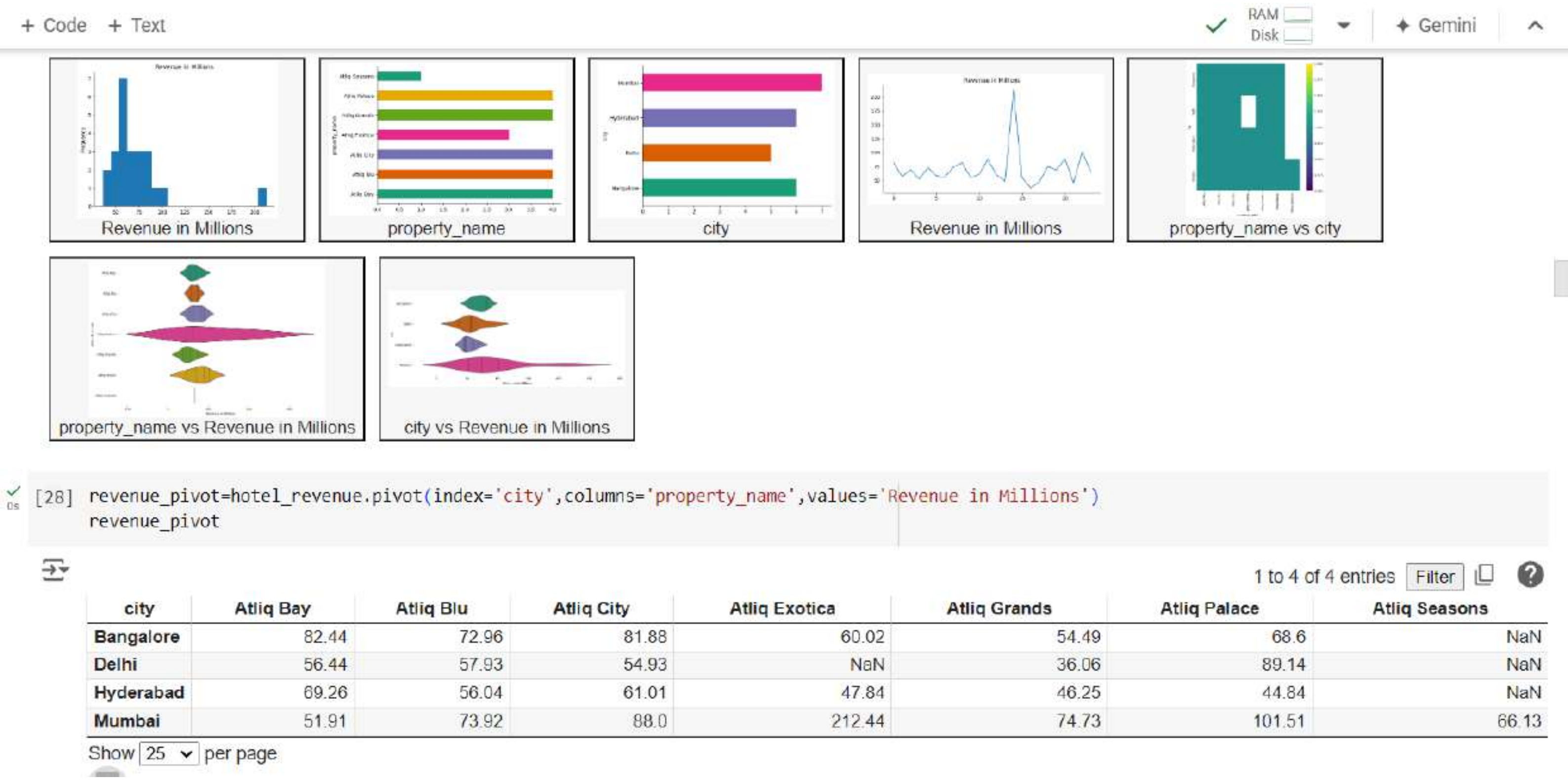


Distributions

Categorical distributions

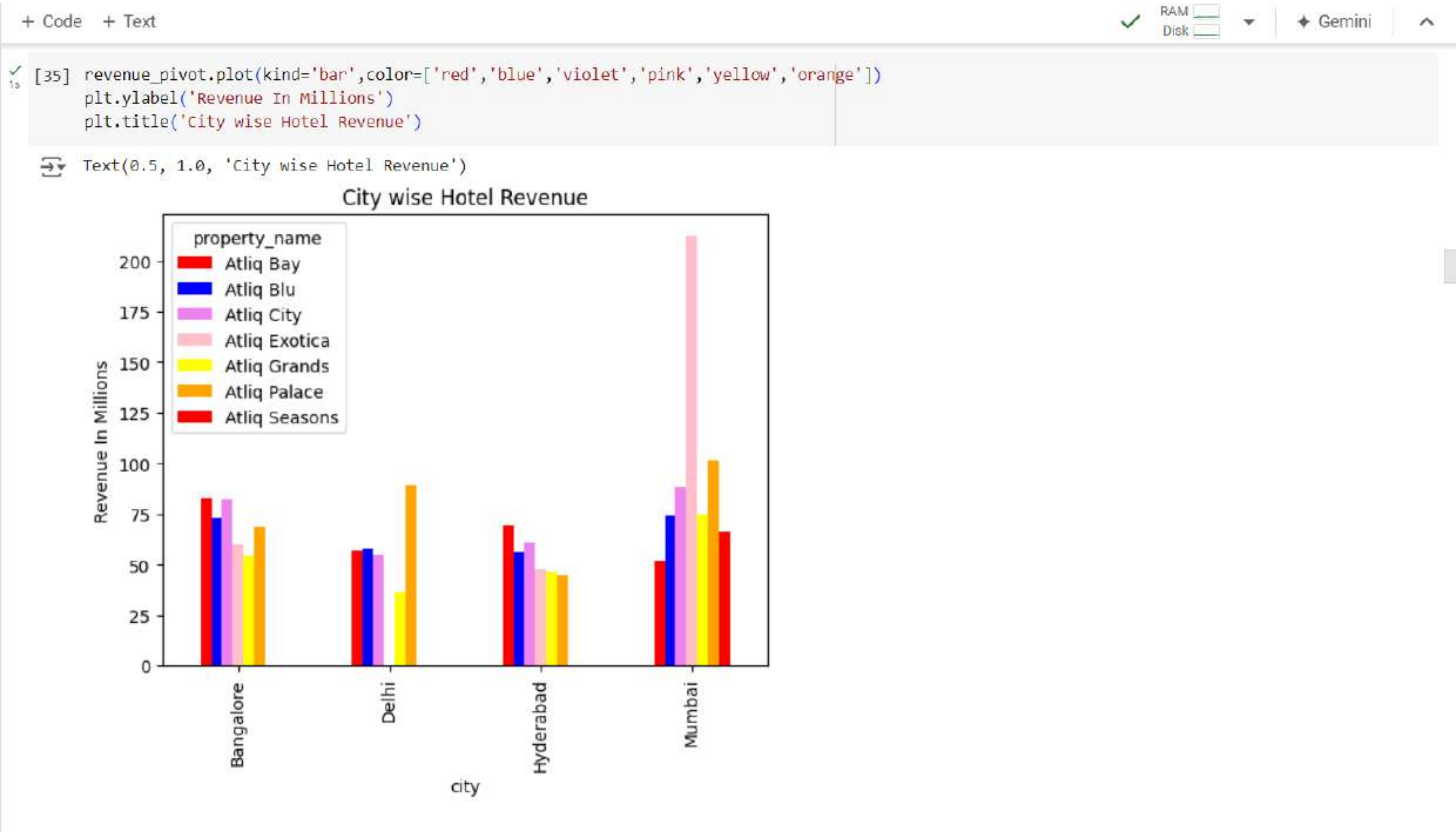
Values

2-d categorical distributions



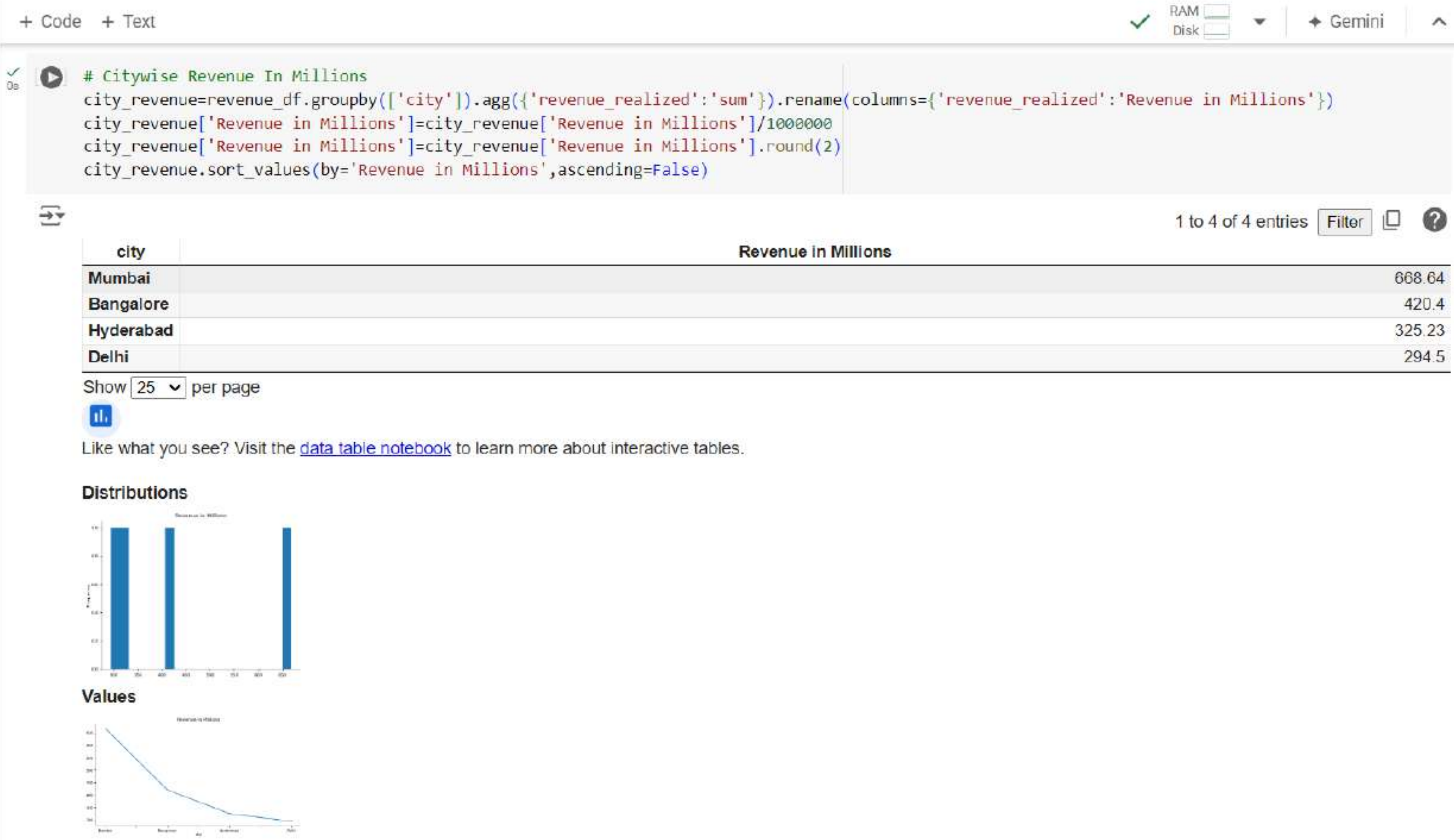
```
[28] revenue_pivot=hotel_revenue.pivot(index='city',columns='property_name',values='Revenue in Millions')
     revenue_pivot
```

city	Atliq Bay	Atliq Blu	Atliq City	Atliq Exotica	Atliq Grands	Atliq Palace	Atliq Seasons
Bangalore	82.44	72.96	81.88	60.02	54.49	68.6	NaN
Delhi	56.44	57.93	54.93	NaN	36.06	89.14	NaN
Hyderabad	69.26	56.04	61.01	47.84	46.25	44.84	NaN
Mumbai	51.91	73.92	88.0	212.44	74.73	101.51	66.13



```
[35] revenue_pivot.plot(kind='bar',color=['red','blue','violet','pink','yellow','orange'])
     plt.ylabel('Revenue In Millions')
     plt.title('City wise Hotel Revenue')
```

```
Text(0.5, 1.0, 'City wise Hotel Revenue')
```



```
# Citywise Revenue In Millions
city_revenue=revenue_df.groupby(['city']).agg({'revenue_realized':'sum'}).rename(columns={'revenue_realized':'Revenue in Millions'})
city_revenue['Revenue in Millions']=city_revenue['Revenue in Millions']/1000000
city_revenue['Revenue in Millions']=city_revenue['Revenue in Millions'].round(2)
city_revenue.sort_values(by='Revenue in Millions',ascending=False)
```

1 to 4 of 4 entries Filter

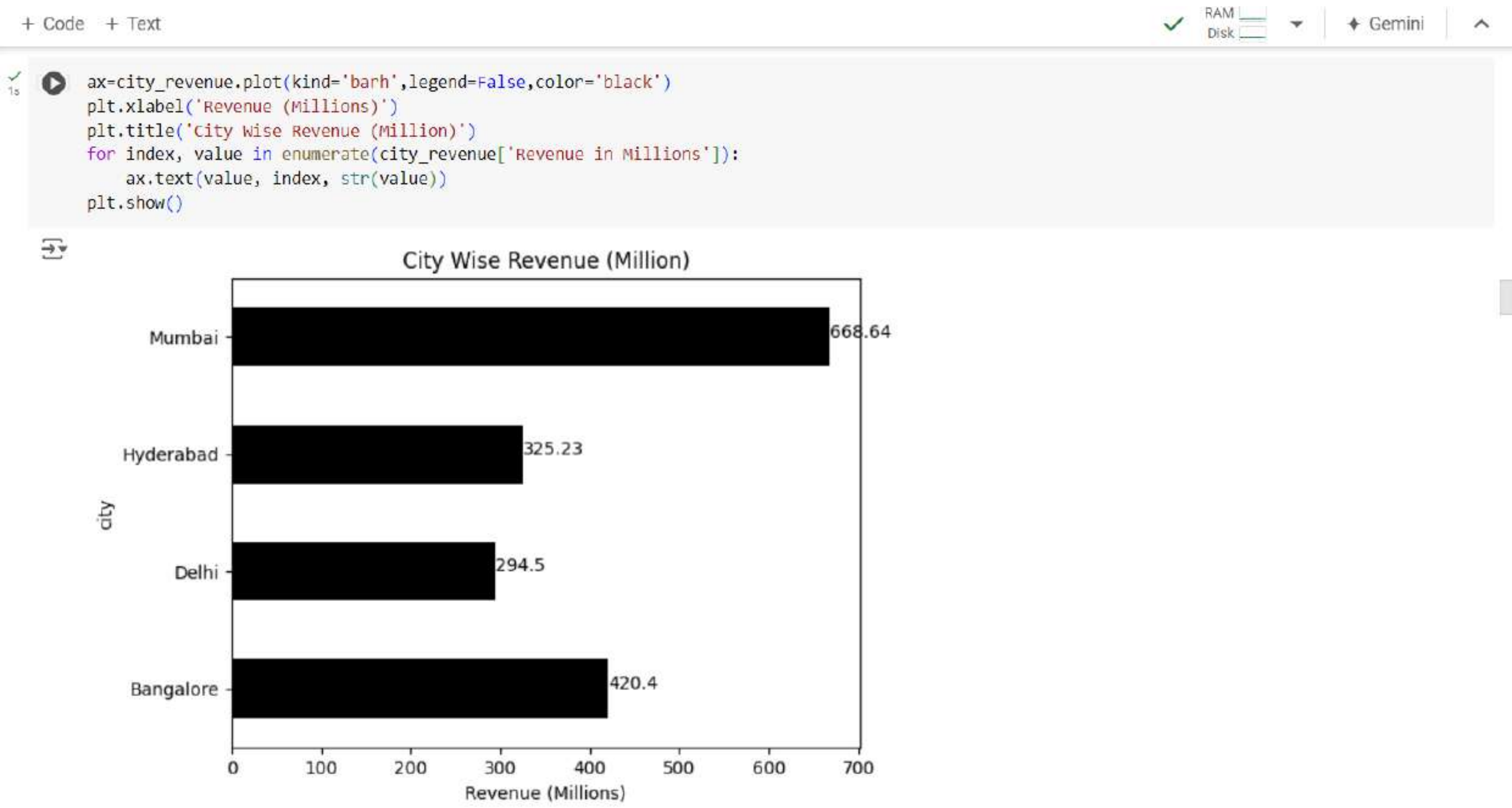
city	Revenue in Millions
Mumbai	668.64
Bangalore	420.4
Hyderabad	325.23
Delhi	294.5

Show 25 per page

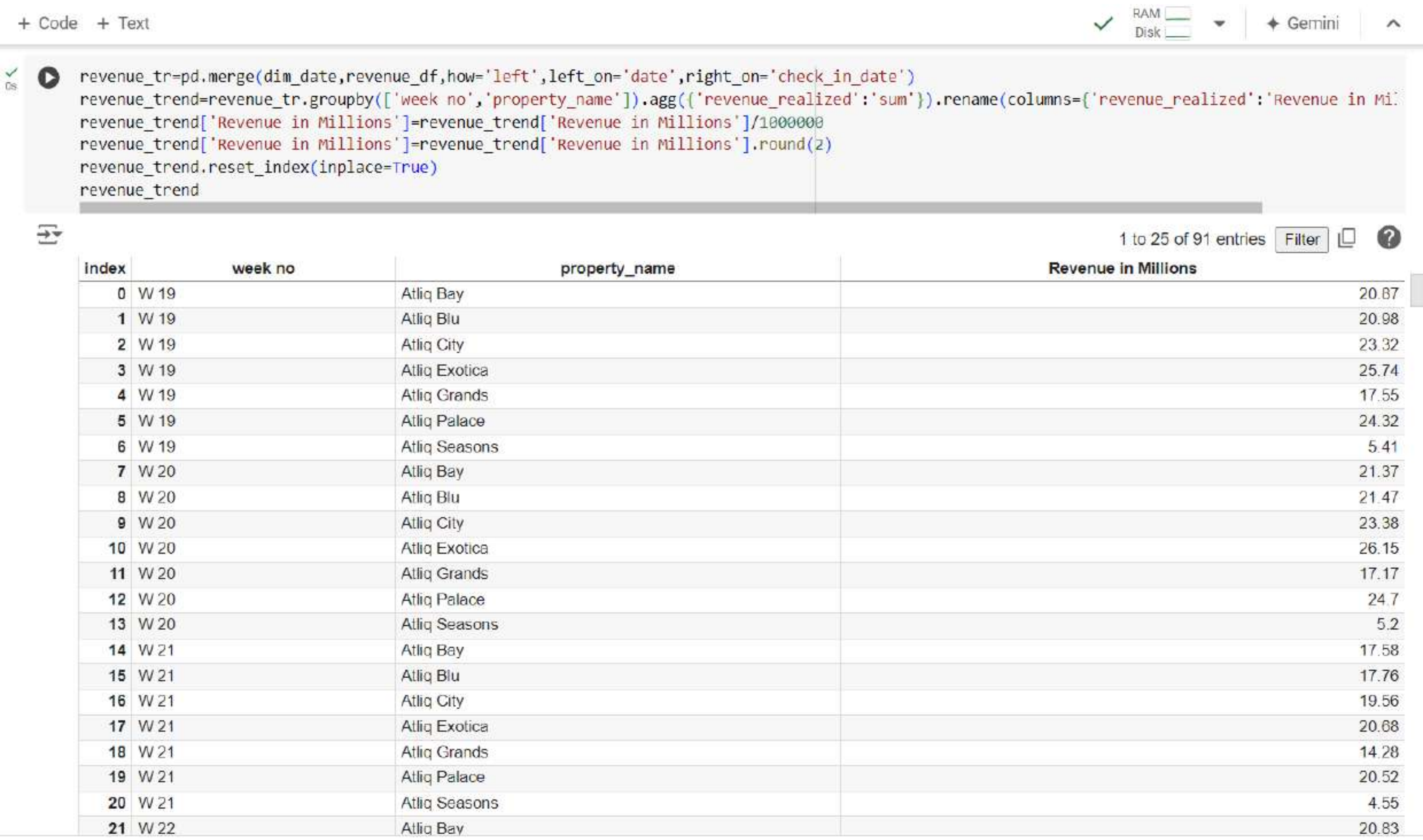
Like what you see? Visit the data table notebook to learn more about interactive tables.

Distributions

Values

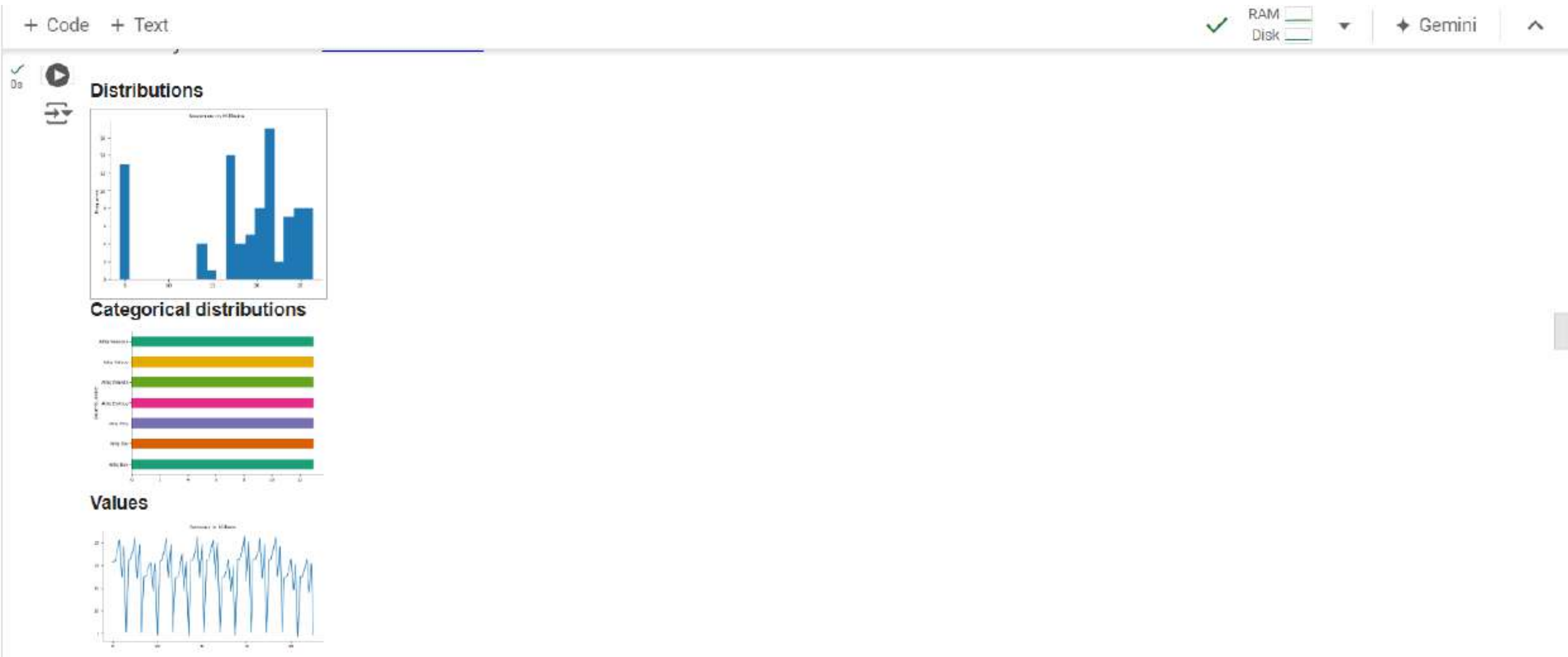


```python
ax=city_revenue.plot(kind='barh',legend=False,color='black')
plt.xlabel('Revenue (Millions)')
plt.title('City Wise Revenue (Million)')
for index, value in enumerate(city_revenue['Revenue in Millions']):
    ax.text(value, index, str(value))
plt.show()
```



```
revenue_tr=pd.merge(dim_date,revenue_df,how='left',left_on='date',right_on='check_in_date')
revenue_trend=revenue_tr.groupby(['week no','property_name']).agg({'revenue_realized':'sum'}).rename(columns={'revenue_realized':'Revenue in Mi
revenue_trend['Revenue in Millions']=revenue_trend['Revenue in Millions']/1000000
revenue_trend['Revenue in Millions']=revenue_trend['Revenue in Millions'].round(2)
revenue_trend.reset_index(inplace=True)
revenue_trend
```

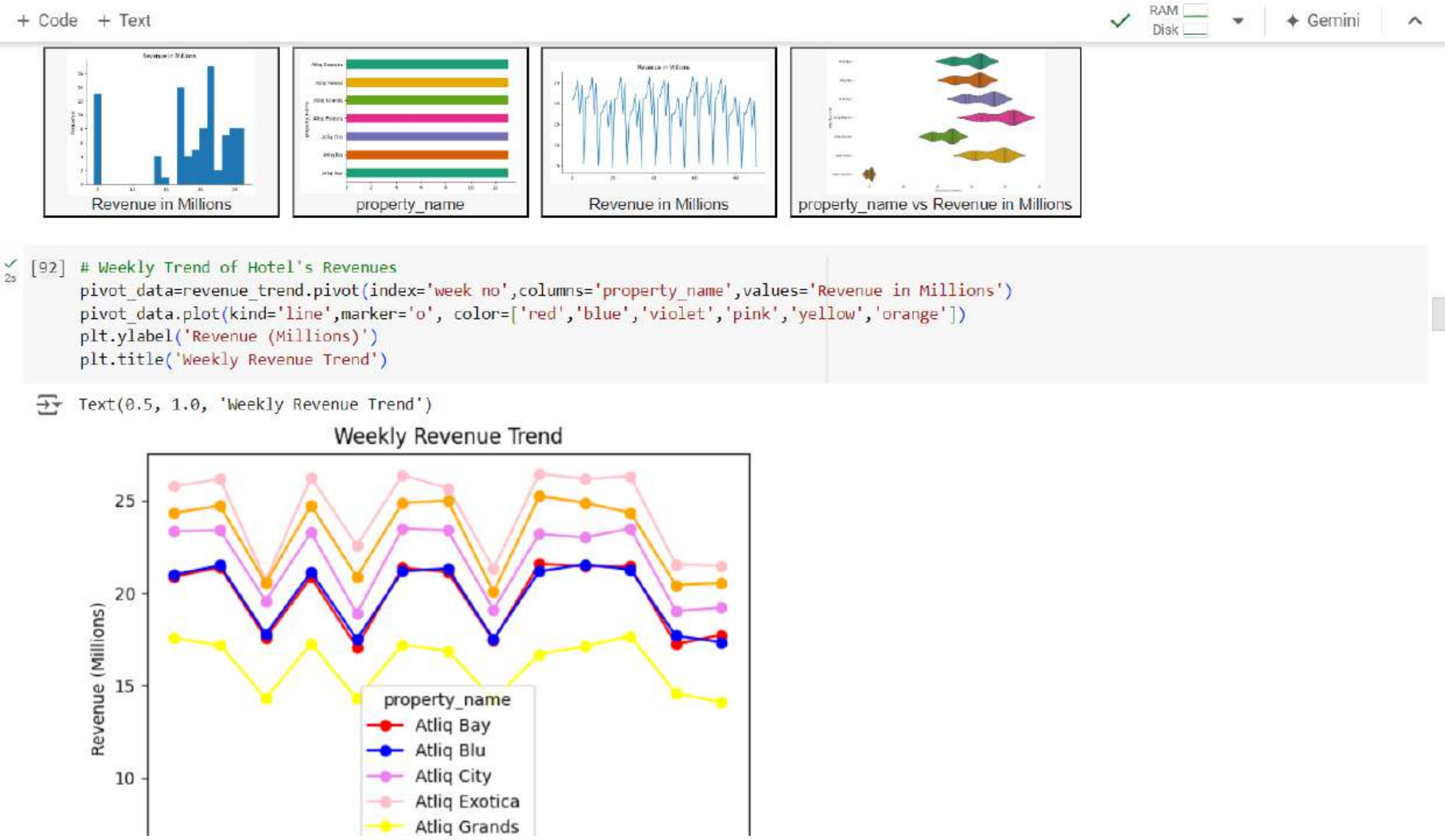
index	week no	property_name	Revenue in Millions
0	W 19	Atliq Bay	20.87
1	W 19	Atliq Blu	20.98
2	W 19	Atliq City	23.32
3	W 19	Atliq Exotica	25.74
4	W 19	Atliq Grands	17.55
5	W 19	Atliq Palace	24.32
6	W 19	Atliq Seasons	5.41
7	W 20	Atliq Bay	21.37
8	W 20	Atliq Blu	21.47
9	W 20	Atliq City	23.38
10	W 20	Atliq Exotica	26.15
11	W 20	Atliq Grands	17.17
12	W 20	Atliq Palace	24.7
13	W 20	Atliq Seasons	5.2
14	W 21	Atliq Bay	17.58
15	W 21	Atliq Blu	17.76
16	W 21	Atliq City	19.56
17	W 21	Atliq Exotica	20.68
18	W 21	Atliq Grands	14.28
19	W 21	Atliq Palace	20.52
20	W 21	Atliq Seasons	4.55
21	W 22	Atliq Bay	20.83



Distributions

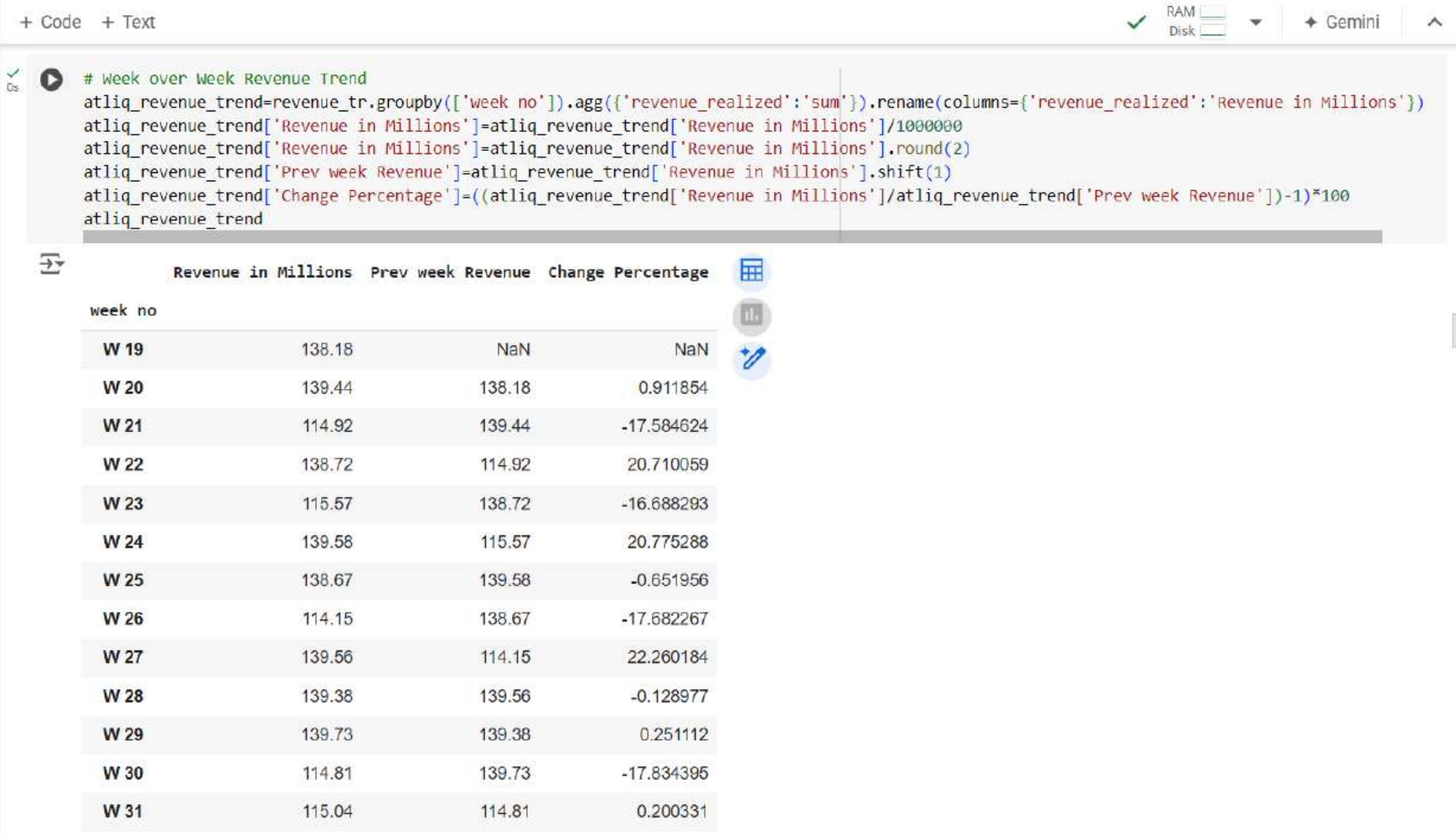
Categorical distributions

Values



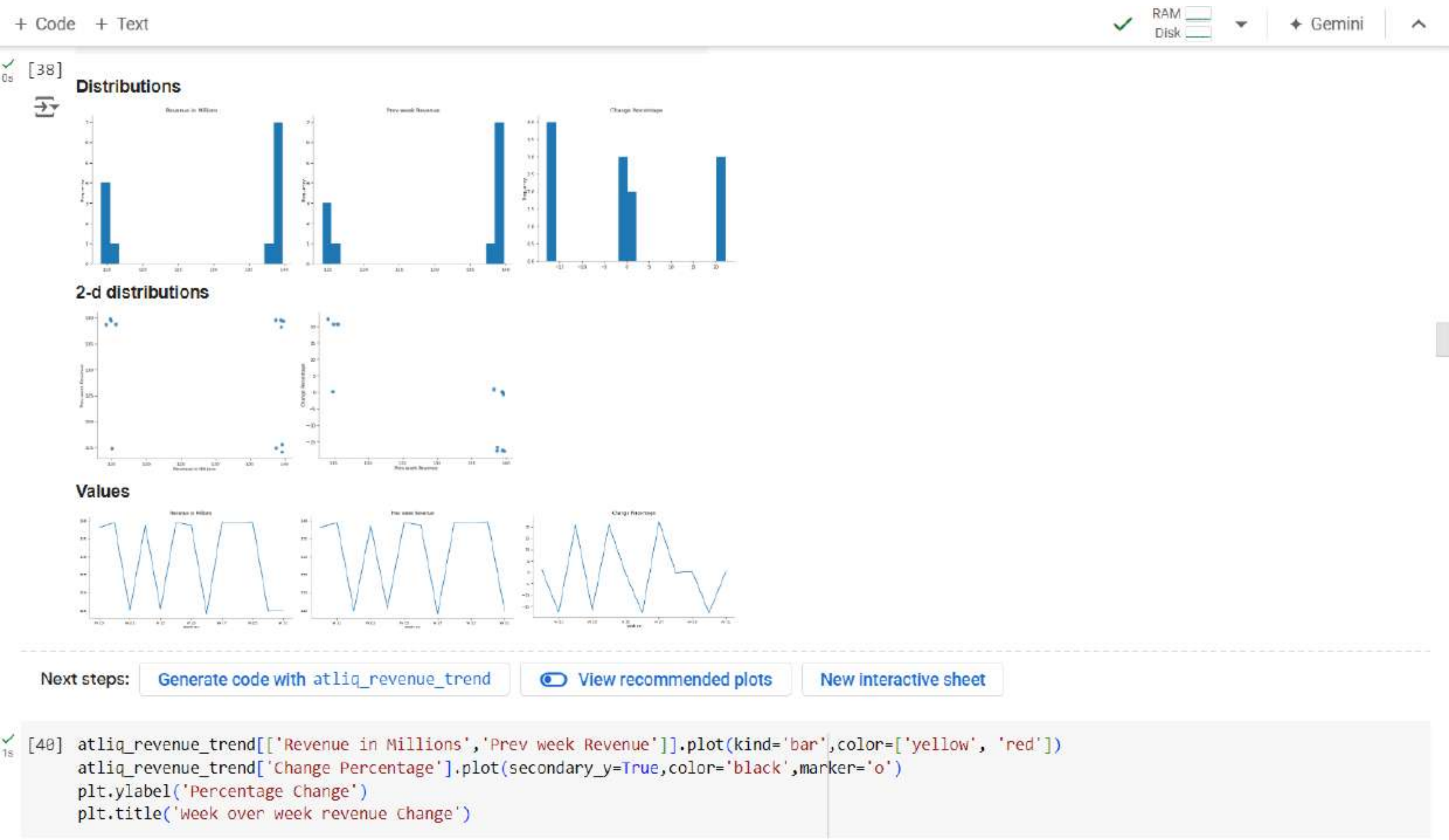
```
[92] # Weekly Trend of Hotel's Revenues
     pivot_data=revenue_trend.pivot(index='week no',columns='property_name',values='Revenue in Millions')
     pivot_data.plot(kind='line',marker='o', color=['red','blue','violet','pink','yellow','orange'])
     plt.ylabel('Revenue (Millions)')
     plt.title('Weekly Revenue Trend')
```

```
Text(0.5, 1.0, 'Weekly Revenue Trend')
```

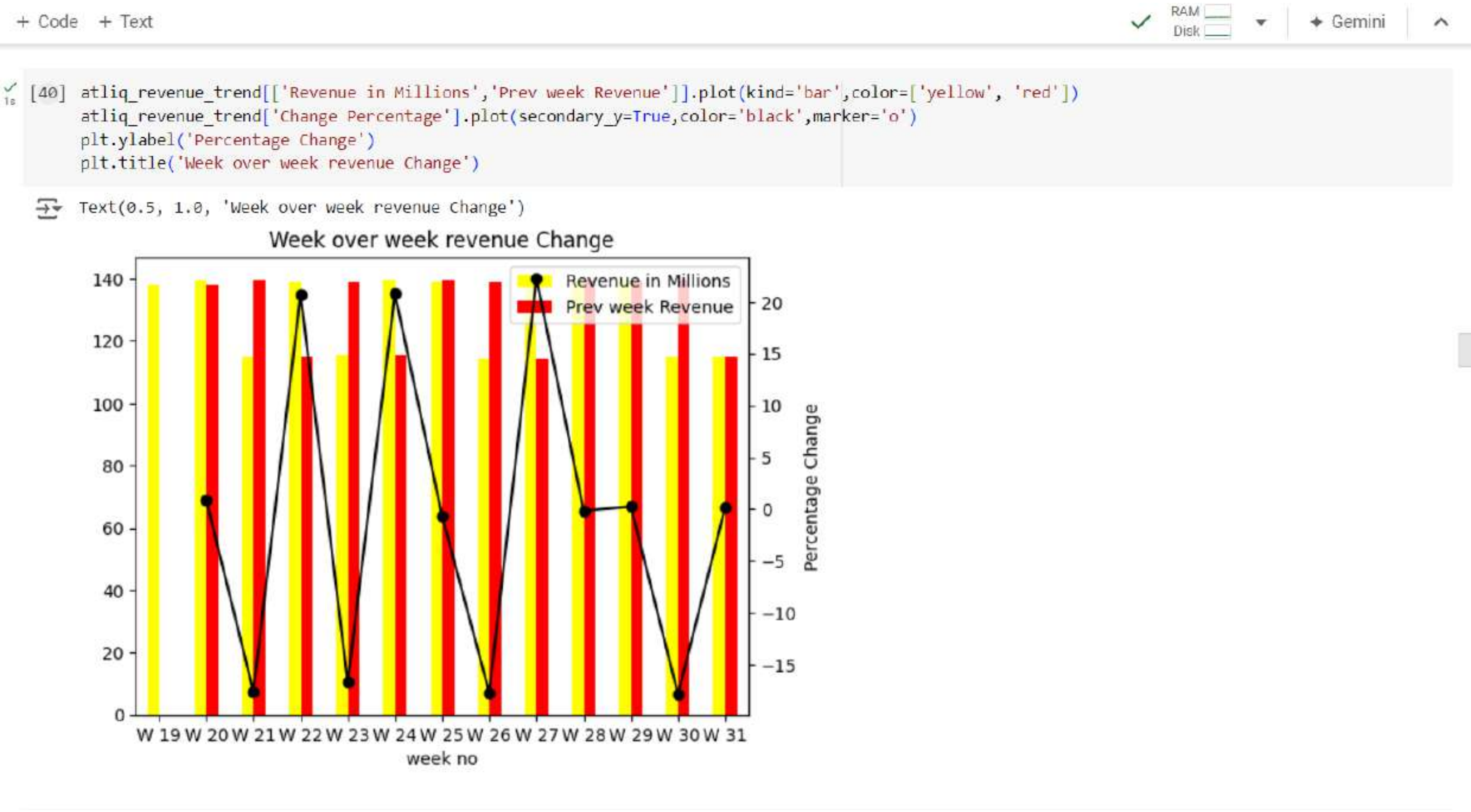


```python
# Week over Week Revenue Trend
atliq_revenue_trend=revenue_tr.groupby(['week no']).agg({'revenue_realized':'sum'}).rename(columns={'revenue_realized':'Revenue in Millions'})
atliq_revenue_trend['Revenue in Millions']=atliq_revenue_trend['Revenue in Millions']/1000000
atliq_revenue_trend['Revenue in Millions']=atliq_revenue_trend['Revenue in Millions'].round(2)
atliq_revenue_trend['Prev week Revenue']=atliq_revenue_trend['Revenue in Millions'].shift(1)
atliq_revenue_trend['Change Percentage']=((atliq_revenue_trend['Revenue in Millions']/atliq_revenue_trend['Prev week Revenue'])-1)*100
atliq_revenue_trend
```

	Revenue in Millions	Prev week Revenue	Change Percentage
week no			
W 19	138.18	NaN	NaN
W 20	139.44	138.18	0.911854
W 21	114.92	139.44	-17.584624
W 22	138.72	114.92	20.710059
W 23	115.57	138.72	-16.688293
W 24	139.58	115.57	20.775288
W 25	138.67	139.58	-0.651956
W 26	114.15	138.67	-17.682267
W 27	139.56	114.15	22.260184
W 28	139.38	139.56	-0.128977
W 29	139.73	139.38	0.251112
W 30	114.81	139.73	-17.834395
W 31	115.04	114.81	0.200331

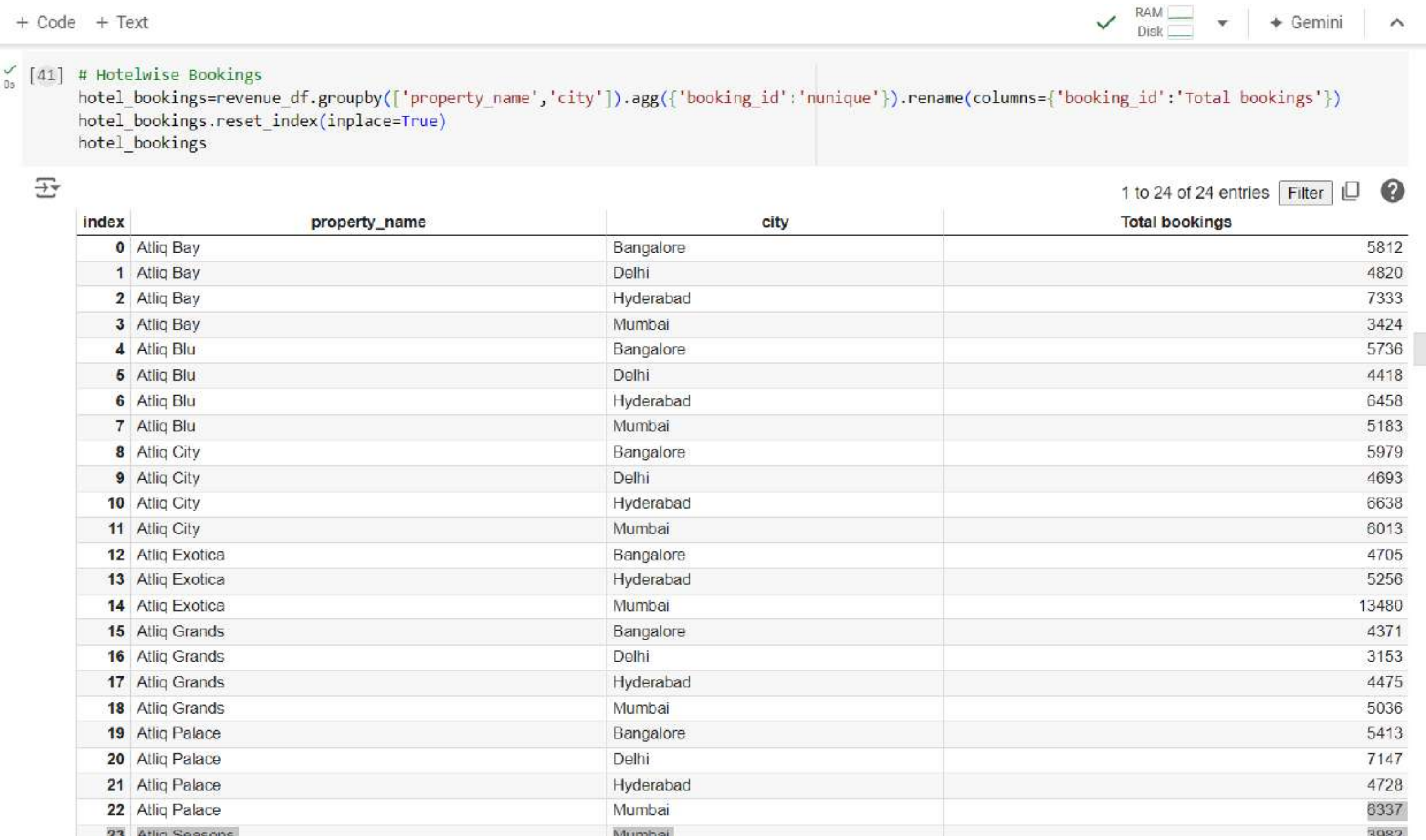


```
atliq_revenue_trend[['Revenue in Millions','Prev week Revenue']].plot(kind='bar',color=['yellow', 'red'])
atliq_revenue_trend['Change Percentage'].plot(secondary_y=True,color='black',marker='o')
plt.ylabel('Percentage Change')
plt.title('Week over week revenue change')
```



```
atliq_revenue_trend[['Revenue in Millions','Prev week Revenue']].plot(kind='bar',color=['yellow', 'red'])
atliq_revenue_trend['Change Percentage'].plot(secondary_y=True,color='black',marker='o')
plt.ylabel('Percentage Change')
plt.title('Week over week revenue Change')
```

```
Text(0.5, 1.0, 'Week over week revenue Change')
```

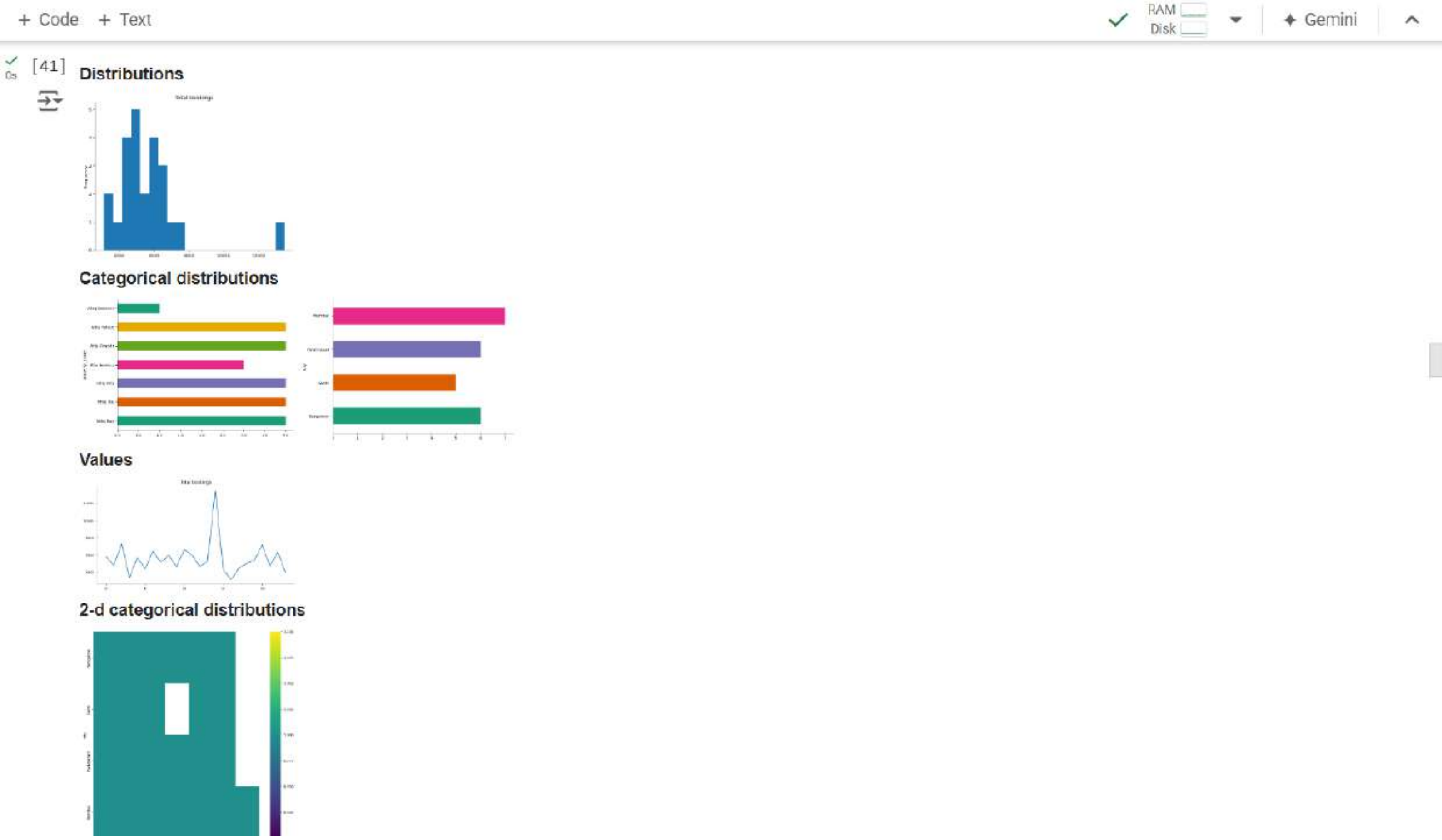


[41]

```
# Hotelwise Bookings
hotel_bookings=revenue_df.groupby(['property_name','city']).agg({'booking_id':'nunique'}).rename(columns={'booking_id':'Total bookings'})
hotel_bookings.reset_index(inplace=True)
hotel_bookings
```

1 to 24 of 24 entries

index	property_name	city	Total bookings
0	Atliq Bay	Bangalore	5812
1	Atliq Bay	Delhi	4820
2	Atliq Bay	Hyderabad	7333
3	Atliq Bay	Mumbai	3424
4	Atliq Blu	Bangalore	5736
5	Atliq Blu	Delhi	4418
6	Atliq Blu	Hyderabad	6458
7	Atliq Blu	Mumbai	5183
8	Atliq City	Bangalore	5979
9	Atliq City	Delhi	4693
10	Atliq City	Hyderabad	6638
11	Atliq City	Mumbai	6013
12	Atliq Exotica	Bangalore	4705
13	Atliq Exotica	Hyderabad	5256
14	Atliq Exotica	Mumbai	13480
15	Atliq Grands	Bangalore	4371
16	Atliq Grands	Delhi	3153
17	Atliq Grands	Hyderabad	4475
18	Atliq Grands	Mumbai	5036
19	Atliq Palace	Bangalore	5413
20	Atliq Palace	Delhi	7147
21	Atliq Palace	Hyderabad	4728
22	Atliq Palace	Mumbai	6337
23	Atliq Seasons	Mumbai	3982

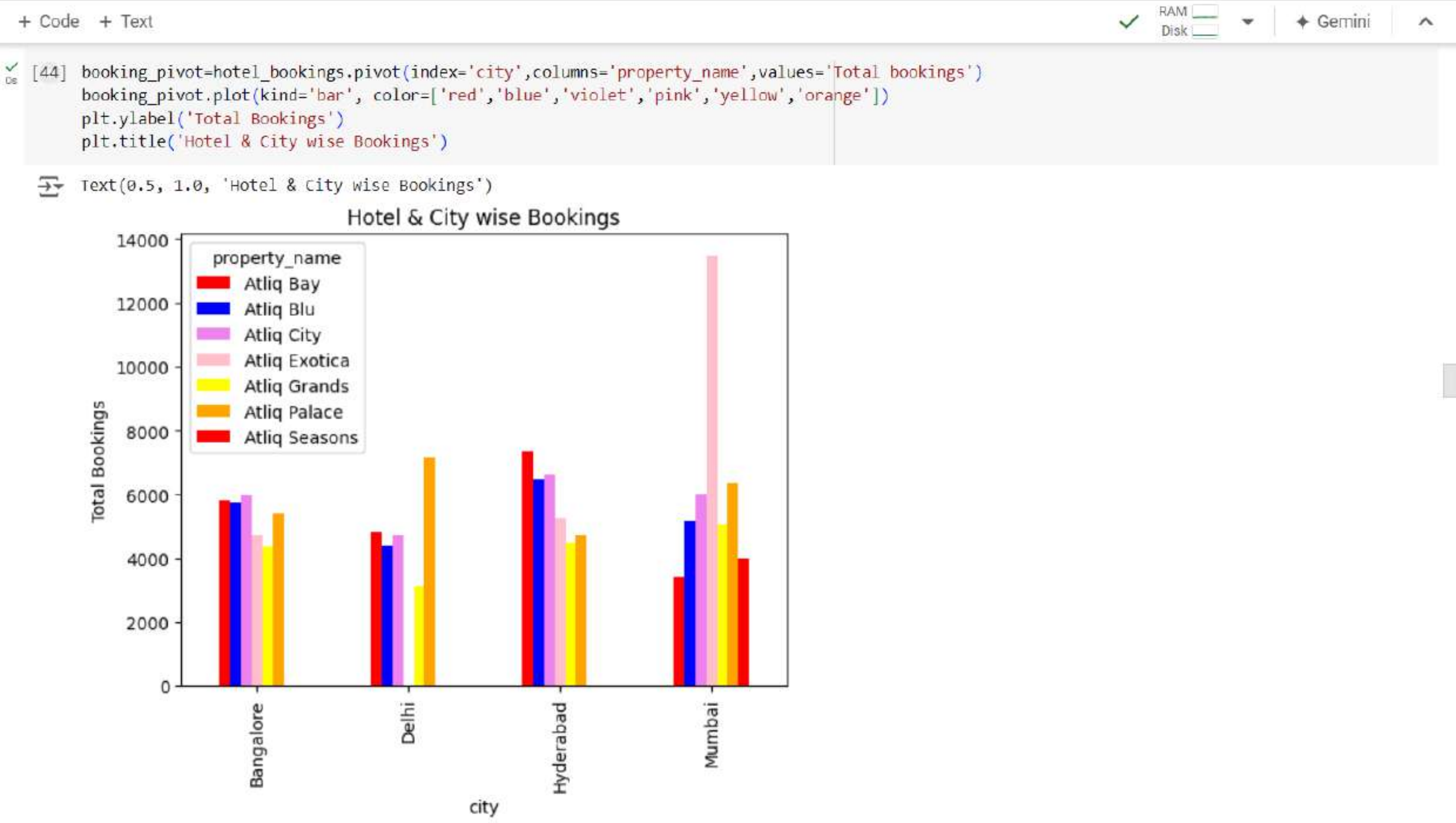


Distributions

Categorical distributions

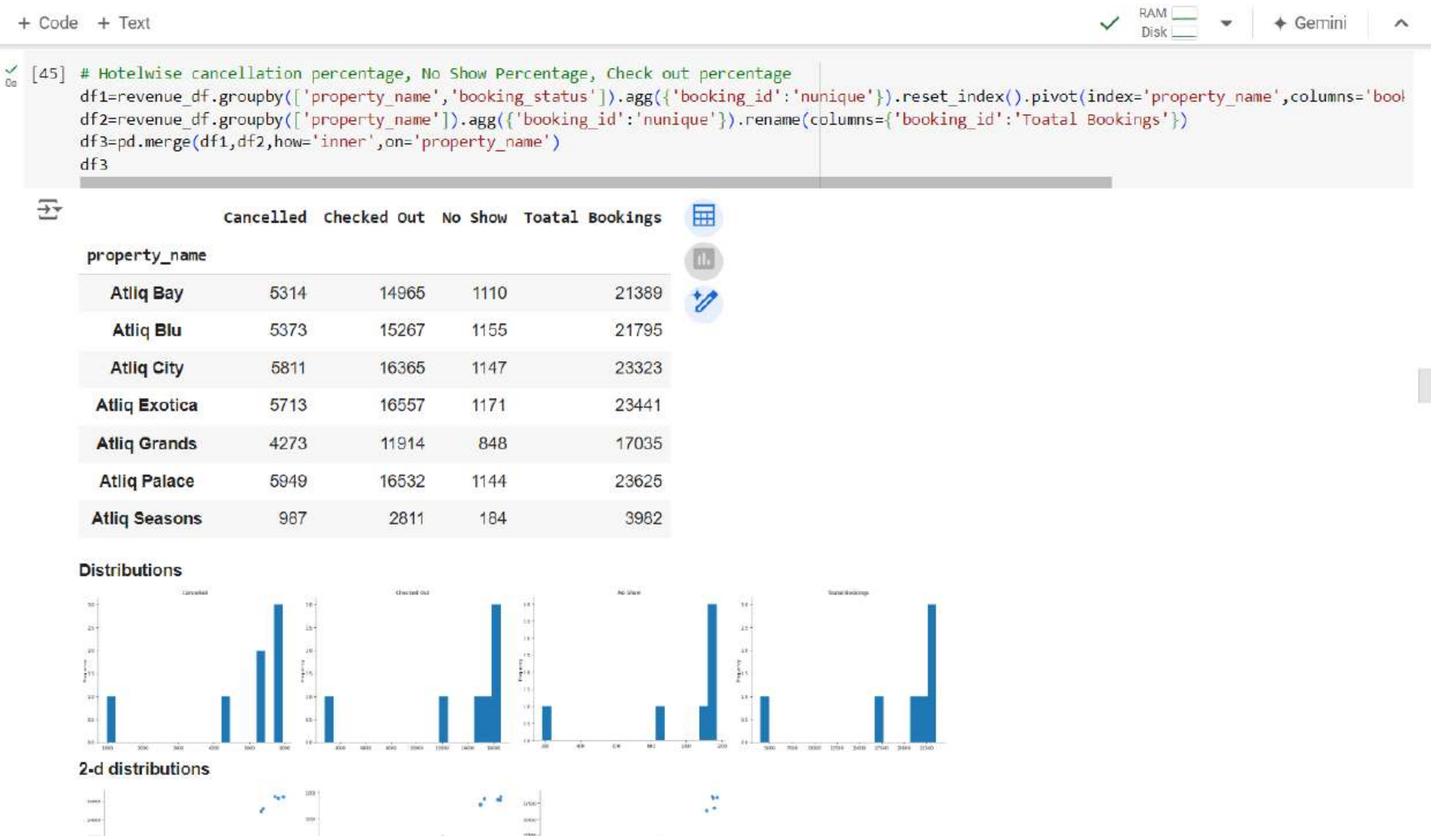
Values

2-d categorical distributions



```
booking_pivot=hotel_bookings.pivot(index='city',columns='property_name',values='Total bookings')
booking_pivot.plot(kind='bar', color=['red','blue','violet','pink','yellow','orange'])
plt.ylabel('Total Bookings')
plt.title('Hotel & City wise Bookings')
```

```
Text(0.5, 1.0, 'Hotel & City wise Bookings')
```

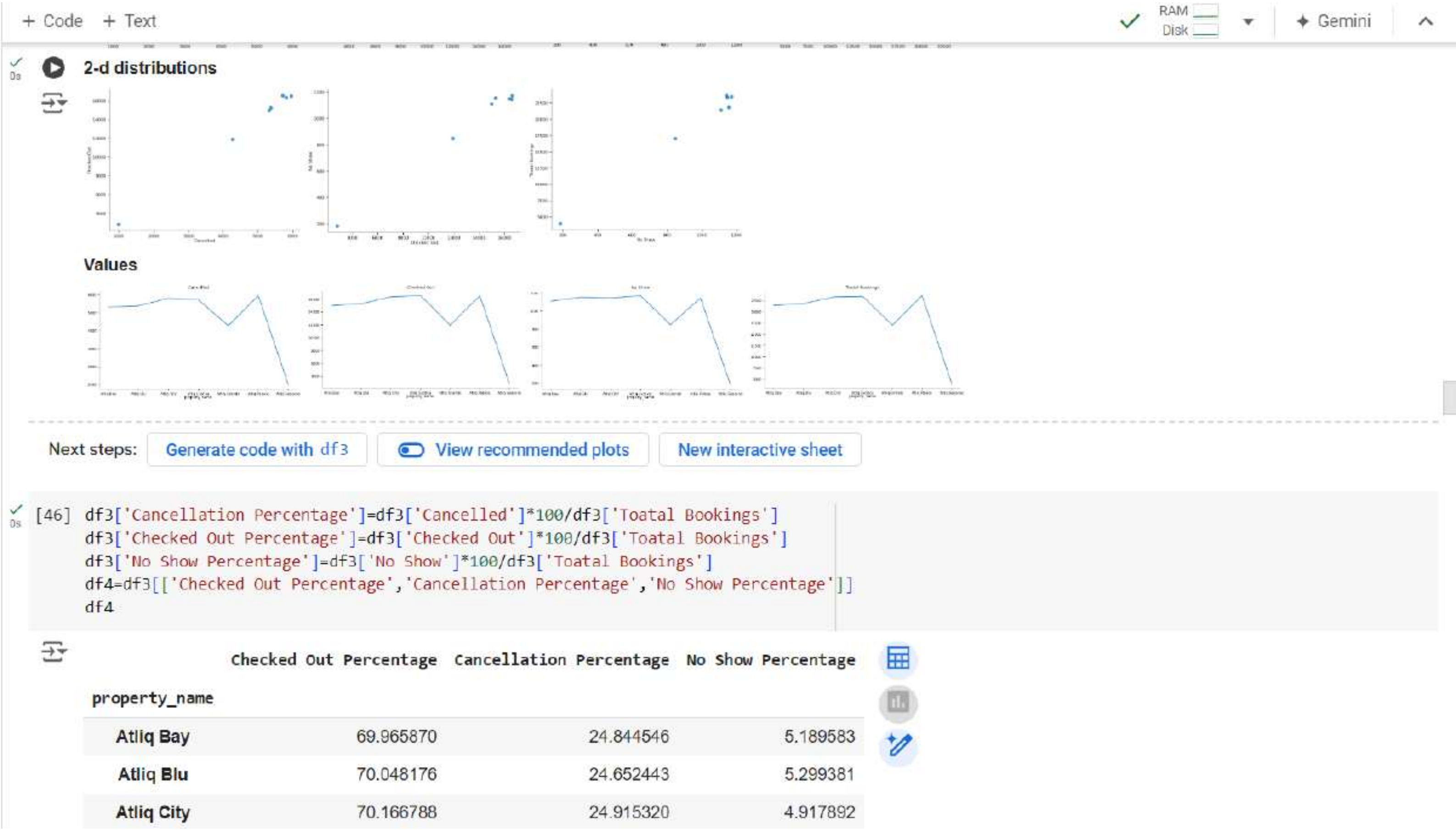


```
[45] # Hotelwise cancellation percentage, No Show Percentage, Check out percentage
df1=revenue_df.groupby(['property_name','booking_status']).agg({'booking_id':'nunique'}).reset_index().pivot(index='property_name',columns='bool
df2=revenue_df.groupby(['property_name']).agg({'booking_id':'nunique'}).rename(columns={'booking_id':'Toatal Bookings'})
df3=pd.merge(df1,df2,how='inner',on='property_name')
df3
```

	Cancelled	Checked Out	No Show	Toatal Bookings
property_name				
Atliq Bay	5314	14965	1110	21389
Atliq Blu	5373	15267	1155	21795
Atliq City	5811	16365	1147	23323
Atliq Exotica	5713	16557	1171	23441
Atliq Grands	4273	11914	848	17035
Atliq Palace	5949	16532	1144	23625
Atliq Seasons	987	2811	184	3982

Distributions

2-d distributions

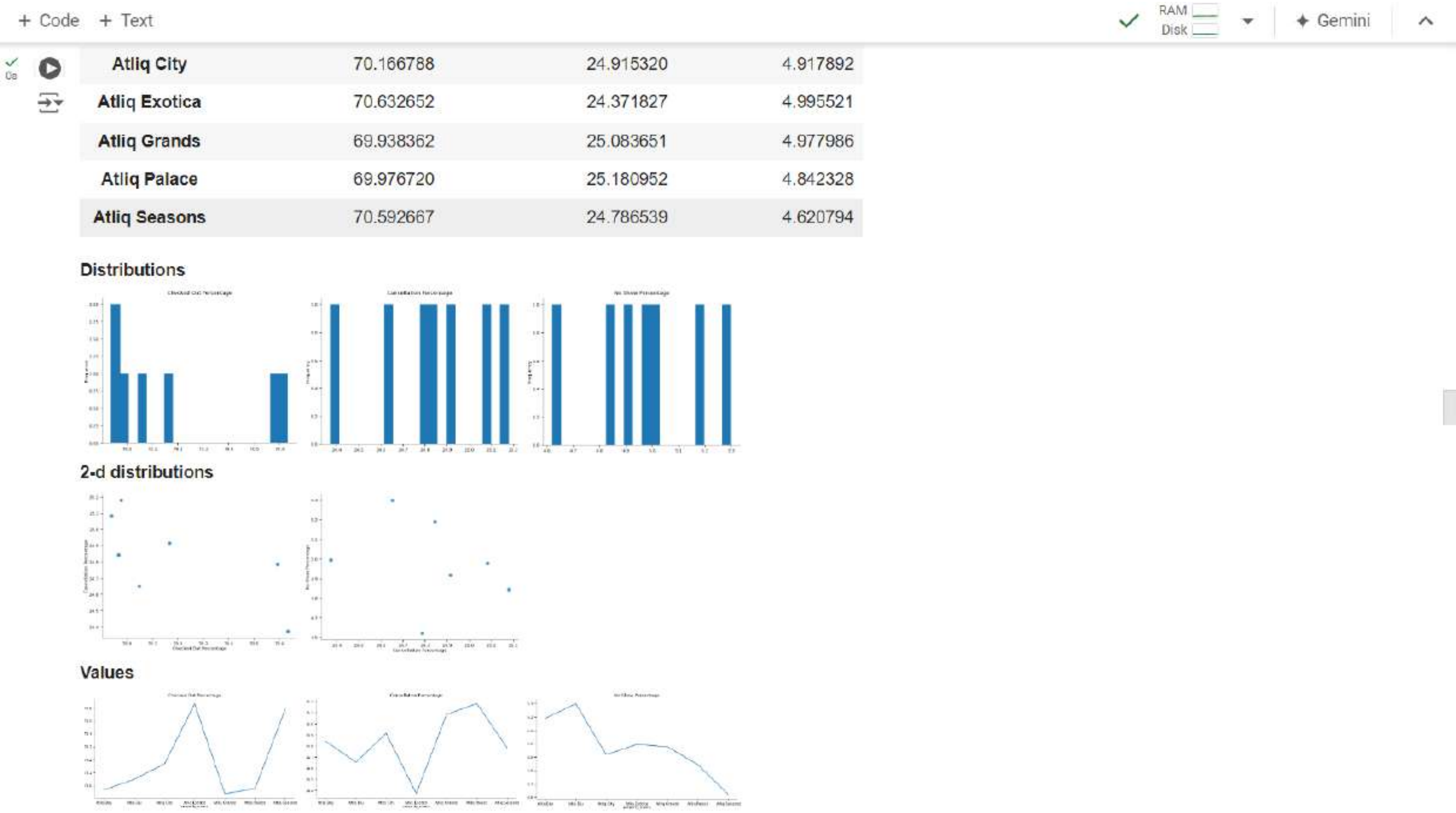


2-d distributions

Values

```
[46] df3['Cancellation Percentage']=df3['Cancelled']*100/df3['Toatal Bookings']
     df3['Checked Out Percentage']=df3['Checked Out']*100/df3['Toatal Bookings']
     df3['No Show Percentage']=df3['No Show']*100/df3['Toatal Bookings']
     df4=df3[['Checked Out Percentage','Cancellation Percentage','No Show Percentage']]
     df4
```

```
               Checked Out Percentage  Cancellation Percentage  No Show Percentage
property_name
   Atliq Bay                69.965870                24.844546            5.189583
   Atliq Blu                70.048176                24.652443            5.299381
   Atliq City               70.166788                24.915320            4.917892
```

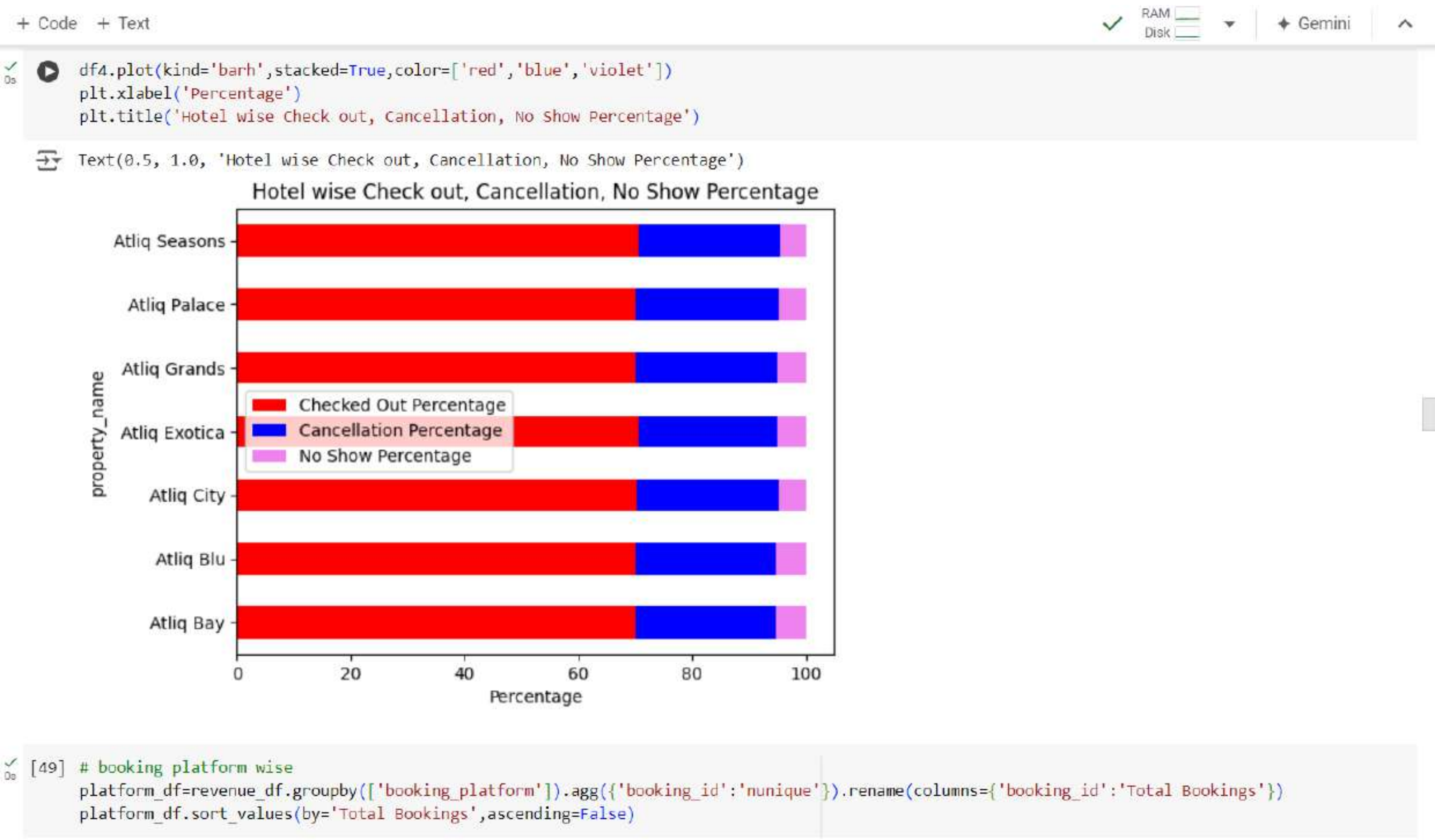


Atliq City	70.166788	24.915320	4.917892
Atliq Exotica	70.632652	24.371827	4.995521
Atliq Grands	69.938362	25.083651	4.977986
Atliq Palace	69.976720	25.180952	4.842328
Atliq Seasons	70.592667	24.786539	4.620794

Distributions

2-d distributions

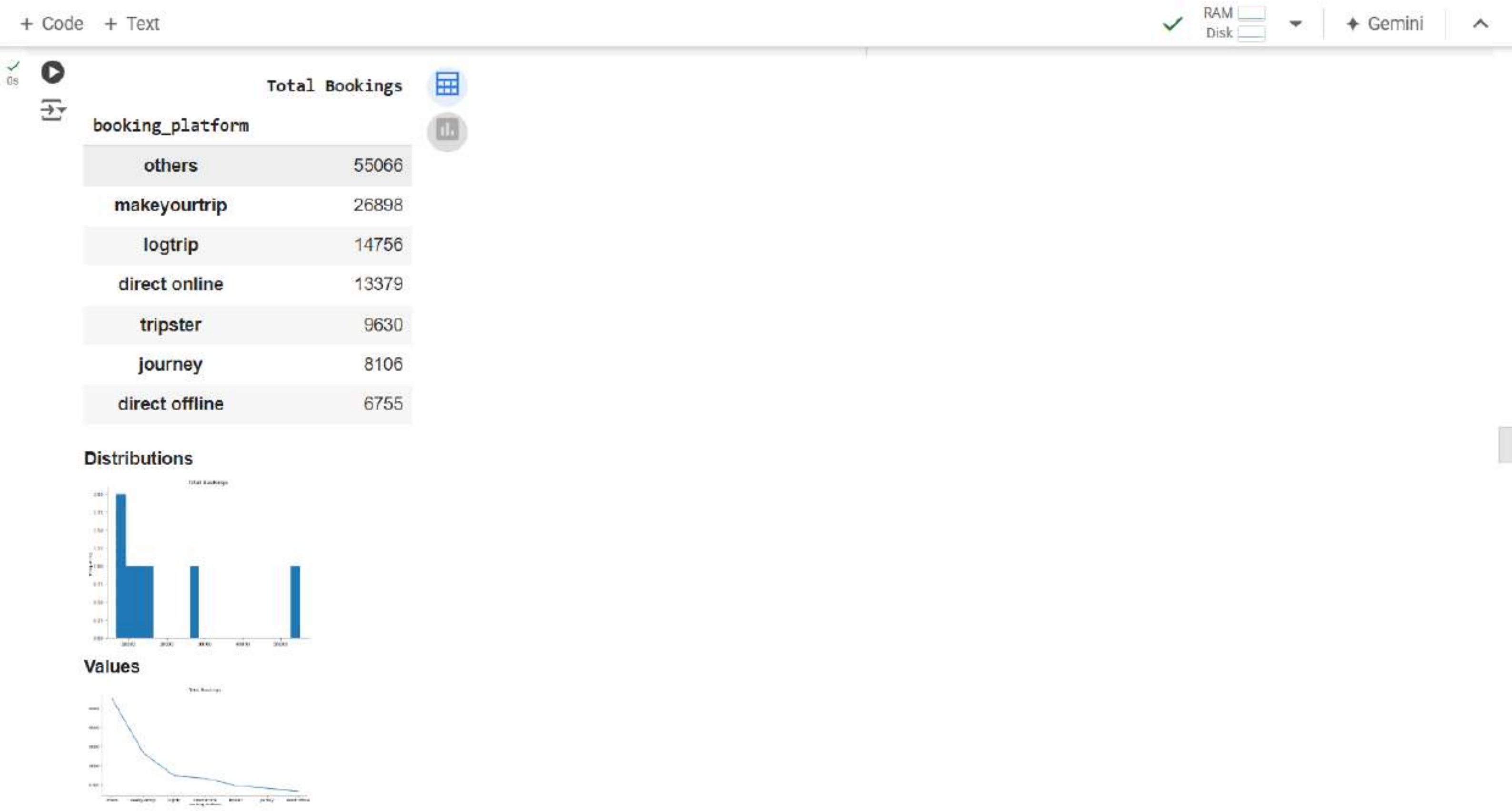
Values



```
df4.plot(kind='barh',stacked=True,color=['red','blue','violet'])
plt.xlabel('Percentage')
plt.title('Hotel wise Check out, Cancellation, No Show Percentage')
```

```
Text(0.5, 1.0, 'Hotel wise Check out, Cancellation, No Show Percentage')
```

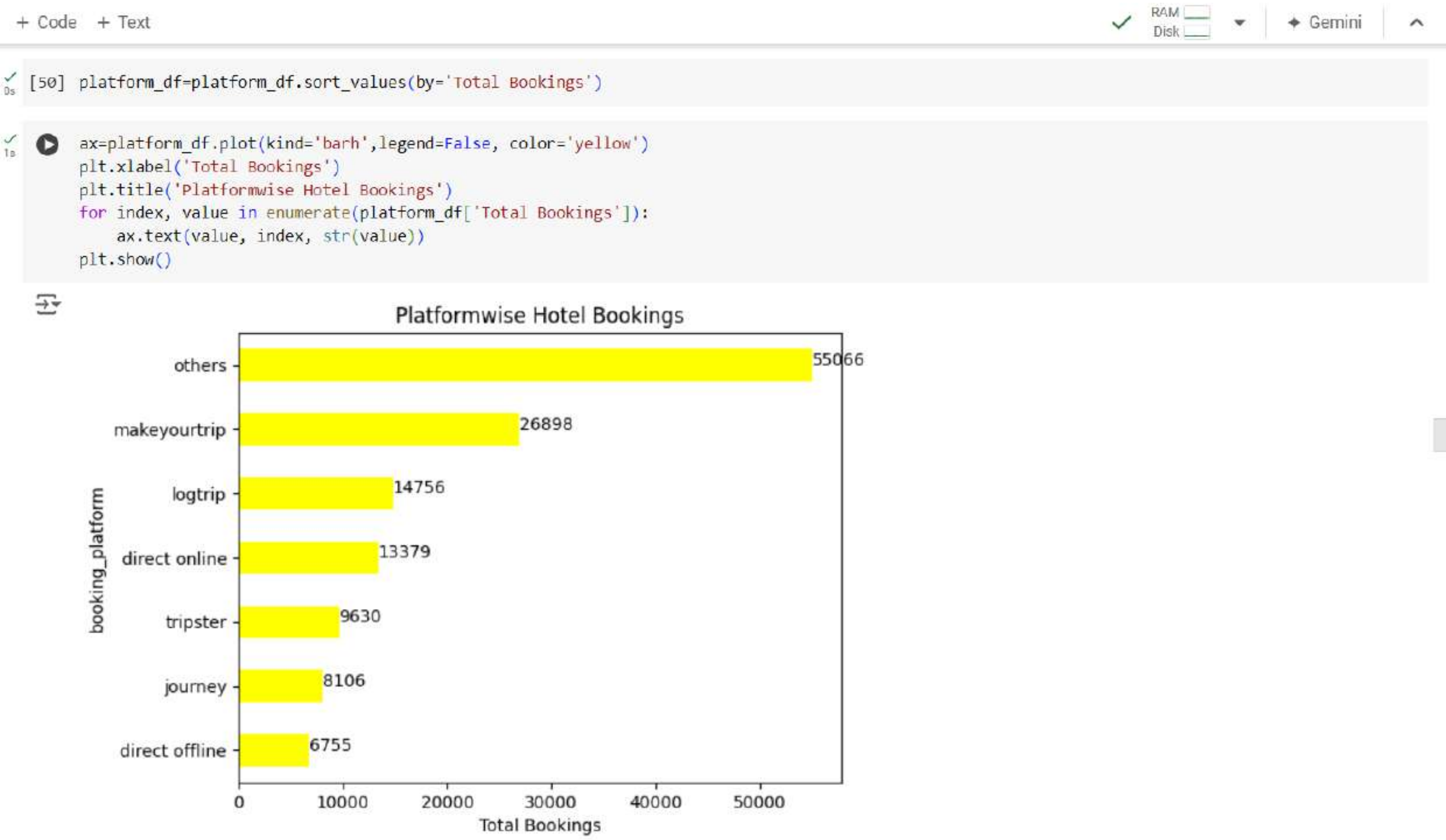
```
[49] # booking platform wise
     platform_df=revenue_df.groupby(['booking_platform']).agg({'booking_id':'nunique'}).rename(columns={'booking_id':'Total Bookings'})
     platform_df.sort_values(by='Total Bookings',ascending=False)
```



booking_platform	Total Bookings
others	55066
makeyourtrip	26898
logtrip	14756
direct online	13379
tripster	9630
journey	8106
direct offline	6755

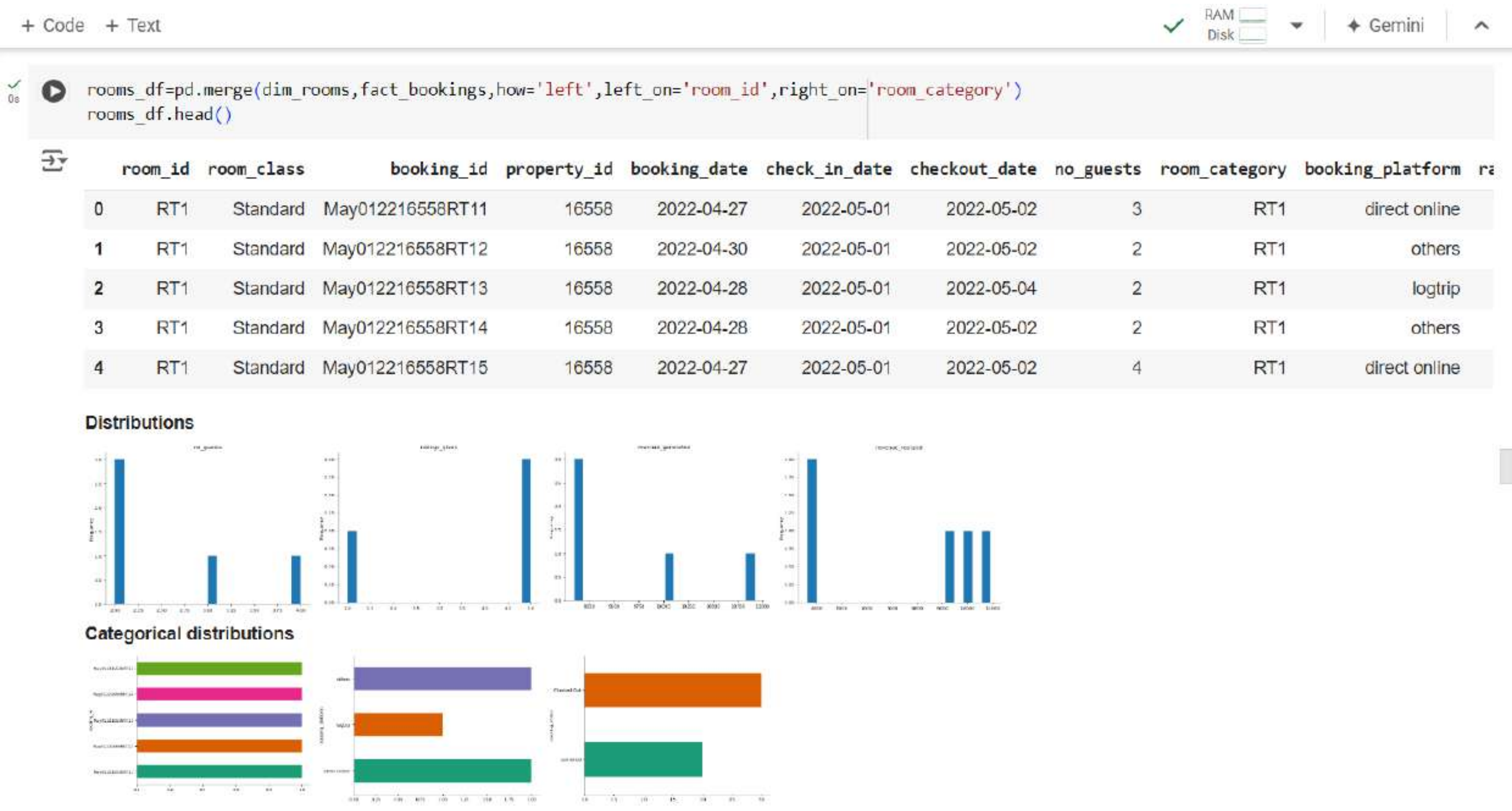
Distributions

Values



```
[50] platform_df=platform_df.sort_values(by='Total Bookings')
```

```
ax=platform_df.plot(kind='barh',legend=False, color='yellow')
plt.xlabel('Total Bookings')
plt.title('Platformwise Hotel Bookings')
for index, value in enumerate(platform_df['Total Bookings']):
    ax.text(value, index, str(value))
plt.show()
```

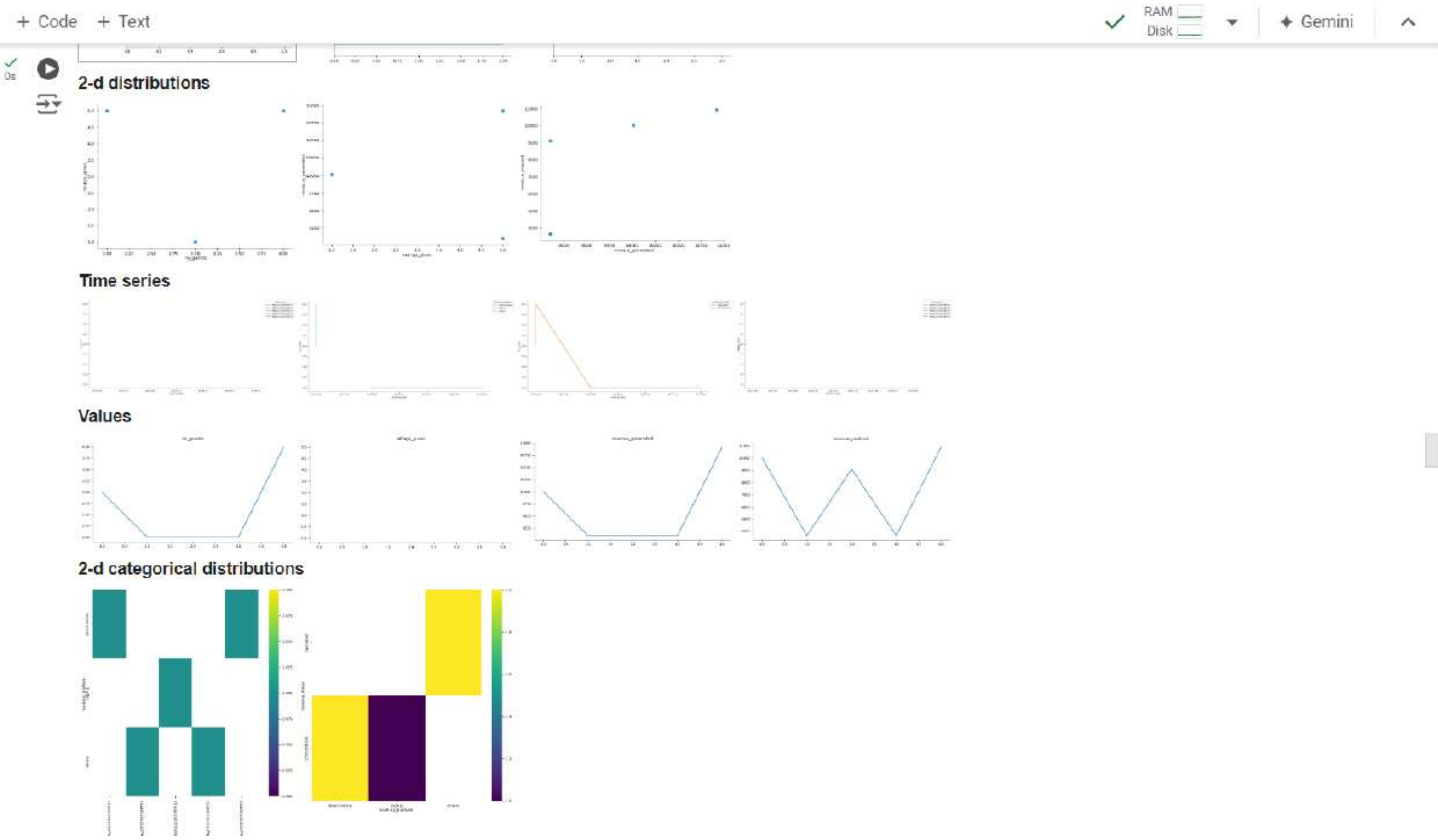


```
rooms_df=pd.merge(dim_rooms,fact_bookings,how='left',left_on='room_id',right_on='room_category')
rooms_df.head()
```

	room_id	room_class	booking_id	property_id	booking_date	check_in_date	checkout_date	no_guests	room_category	booking_platform	ra
0	RT1	Standard	May012216558RT11	16558	2022-04-27	2022-05-01	2022-05-02	3	RT1	direct online	
1	RT1	Standard	May012216558RT12	16558	2022-04-30	2022-05-01	2022-05-02	2	RT1	others	
2	RT1	Standard	May012216558RT13	16558	2022-04-28	2022-05-01	2022-05-04	2	RT1	logtrip	
3	RT1	Standard	May012216558RT14	16558	2022-04-28	2022-05-01	2022-05-02	2	RT1	others	
4	RT1	Standard	May012216558RT15	16558	2022-04-27	2022-05-01	2022-05-02	4	RT1	direct online	

Distributions

Categorical distributions

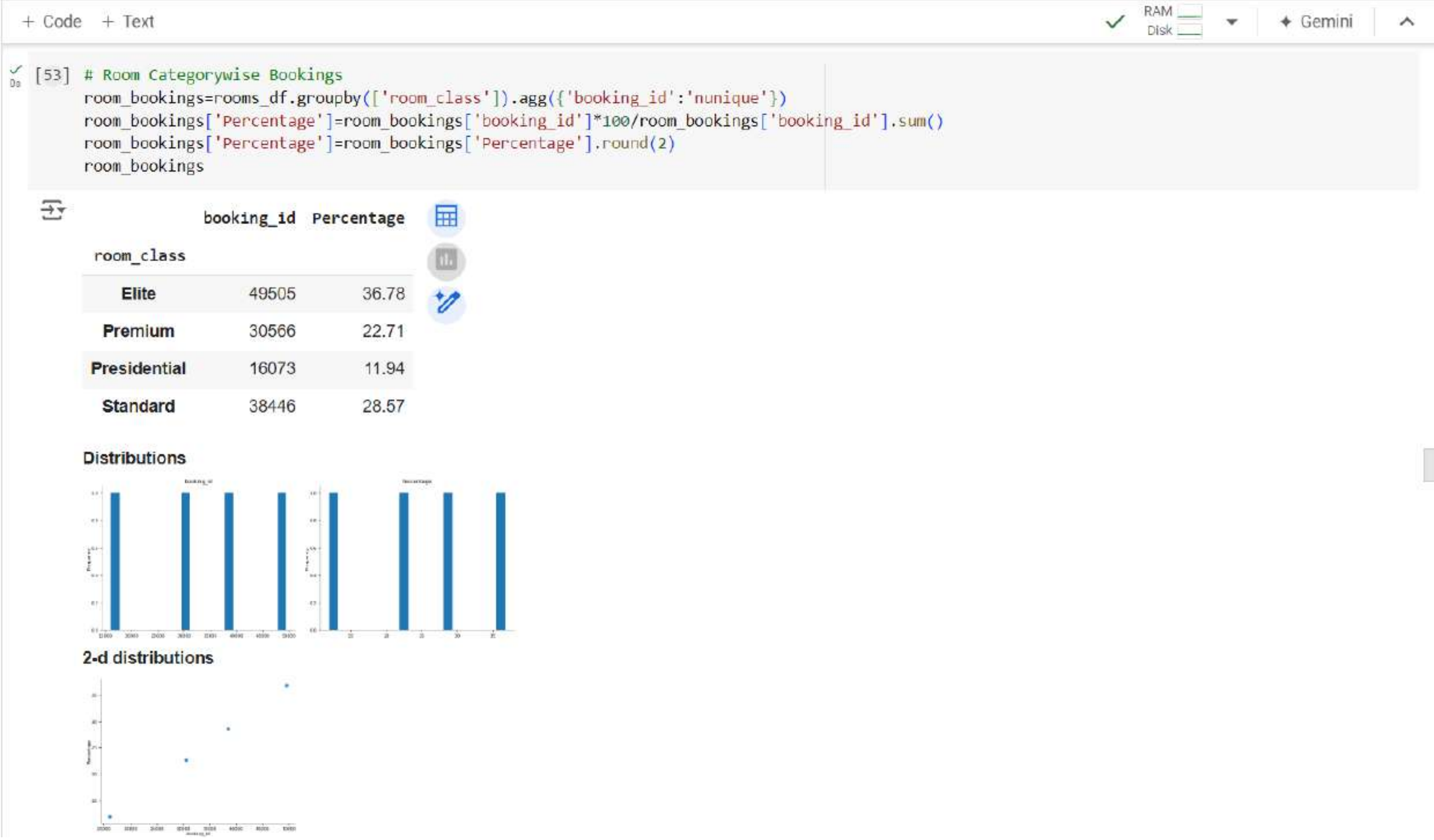


2-d distributions

Time series

Values

2-d categorical distributions

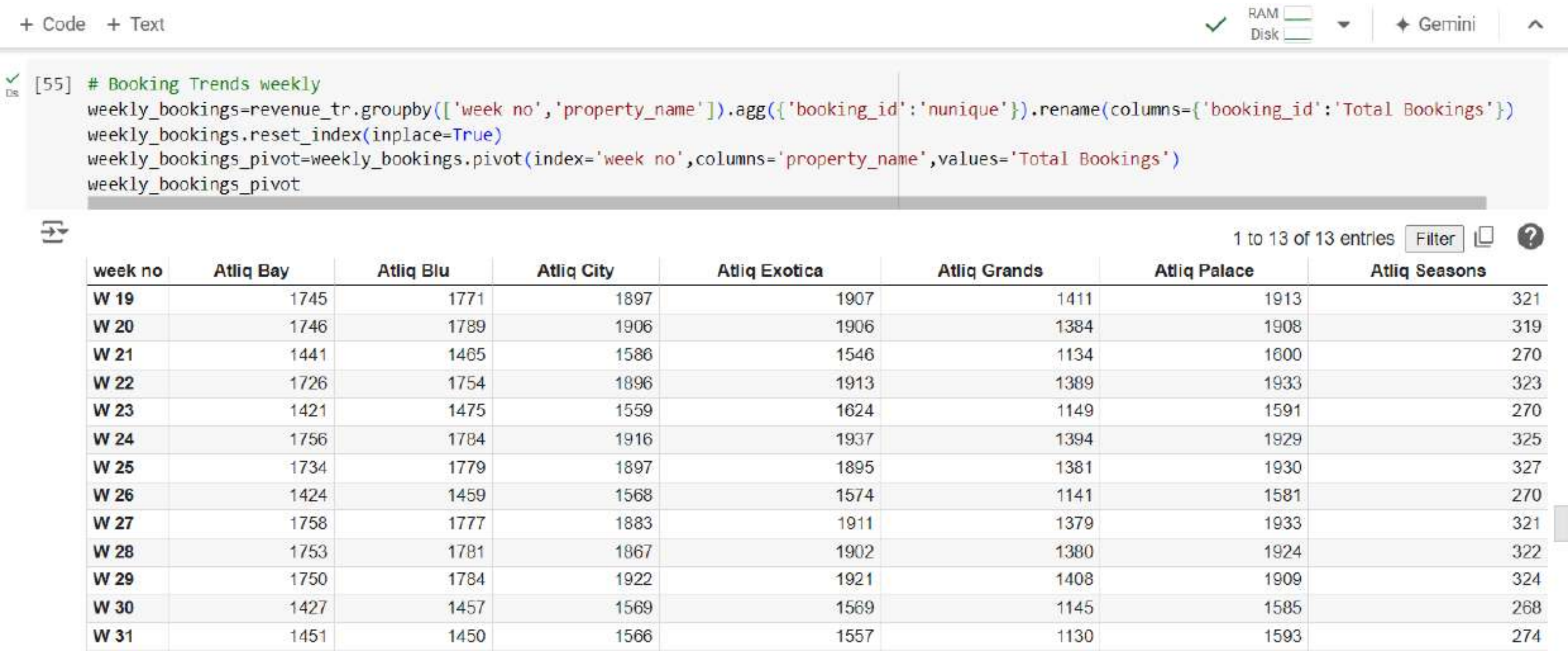


```
# Room Categorywise Bookings
room_bookings=rooms_df.groupby(['room_class']).agg({'booking_id':'nunique'})
room_bookings['Percentage']=room_bookings['booking_id']*100/room_bookings['booking_id'].sum()
room_bookings['Percentage']=room_bookings['Percentage'].round(2)
room_bookings
```

	booking_id	Percentage
room_class		
Elite	49505	36.78
Premium	30566	22.71
Presidential	16073	11.94
Standard	38446	28.57

Distributions

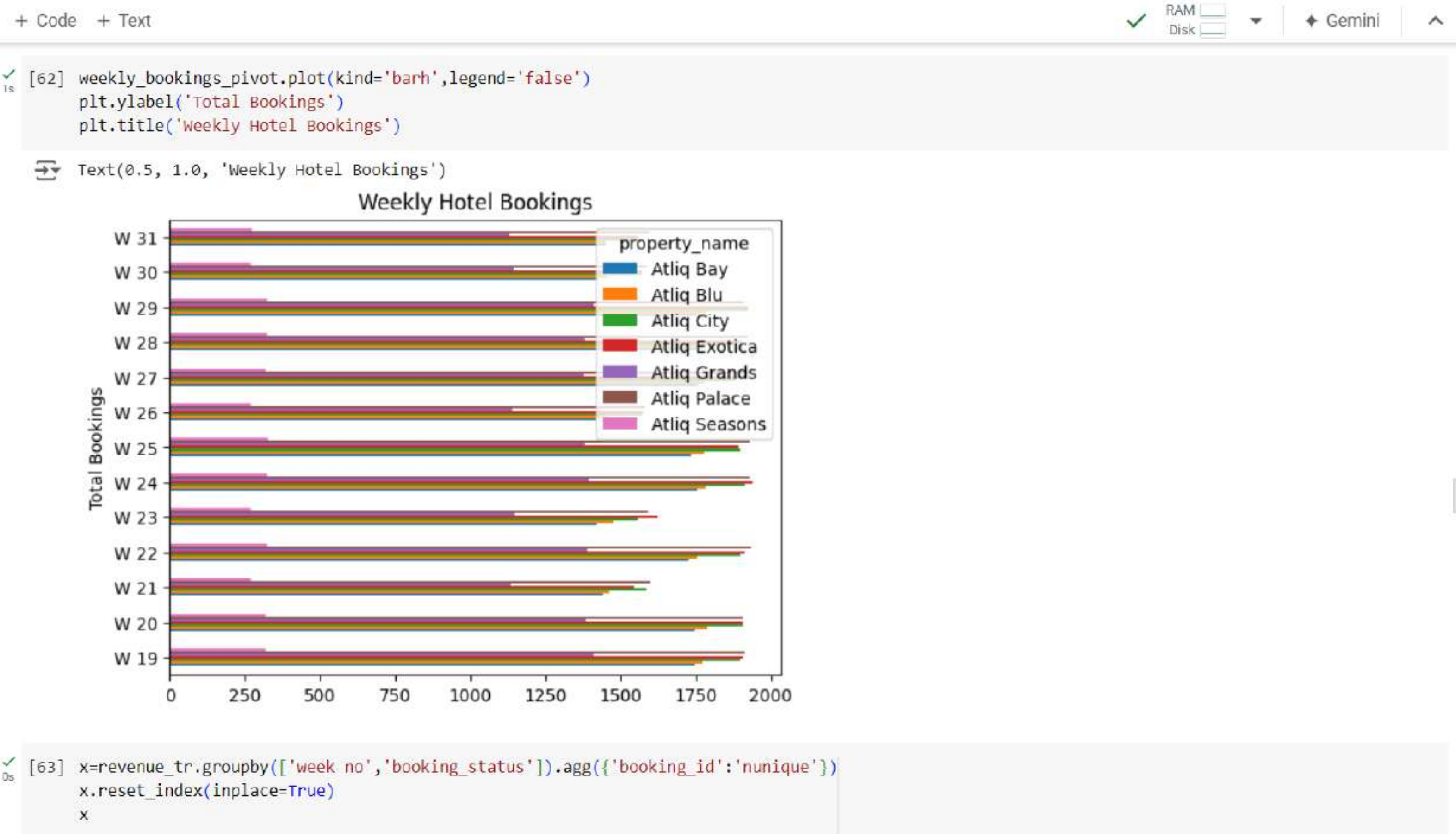
2-d distributions



```
[55] # Booking Trends weekly
     weekly_bookings=revenue_tr.groupby(['week no','property_name']).agg({'booking_id':'nunique'}).rename(columns={'booking_id':'Total Bookings'})
     weekly_bookings.reset_index(inplace=True)
     weekly_bookings_pivot=weekly_bookings.pivot(index='week no',columns='property_name',values='Total Bookings')
     weekly_bookings_pivot
```

1 to 13 of 13 entries Filter

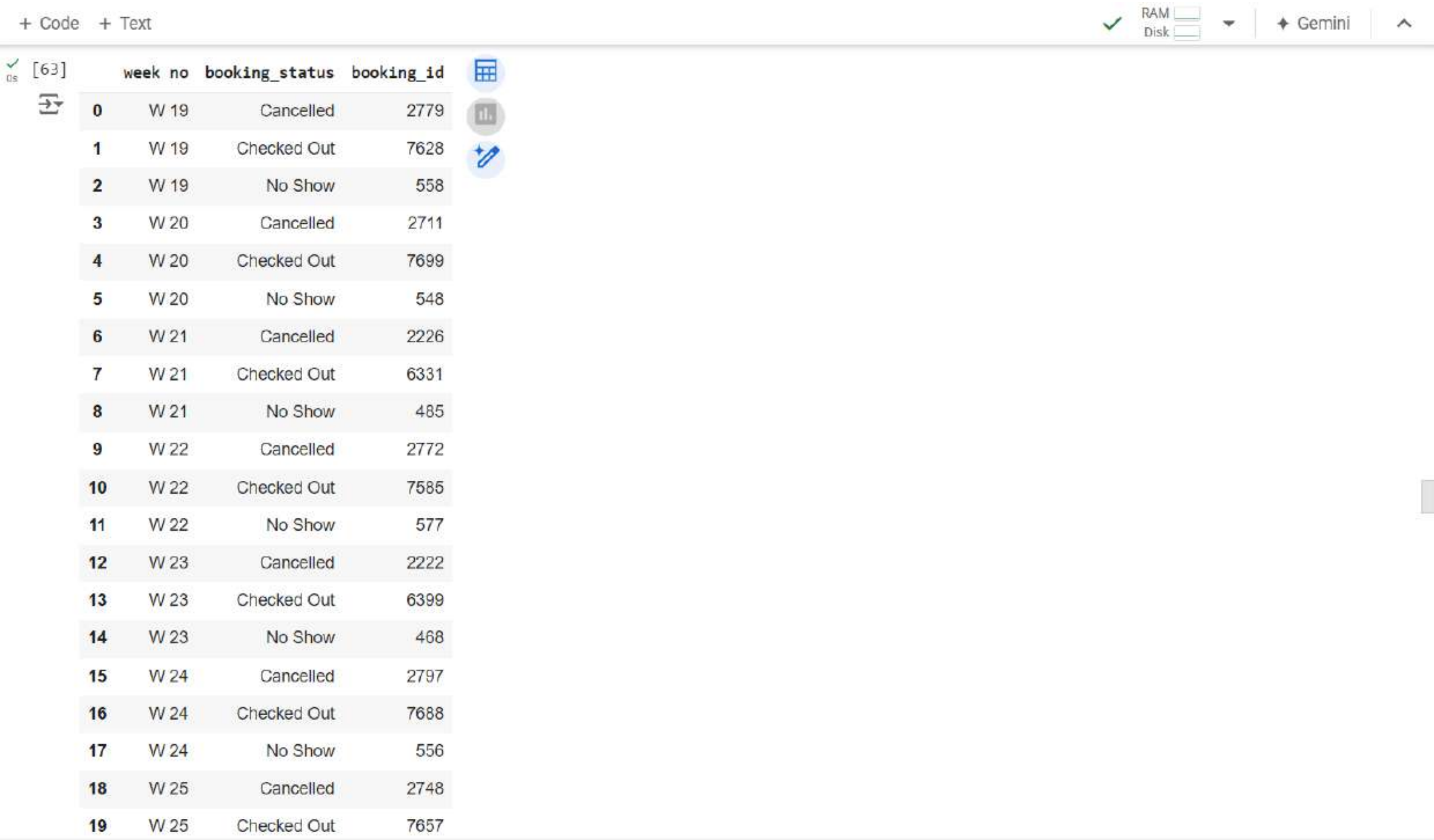
week no	Atliq Bay	Atliq Blu	Atliq City	Atliq Exotica	Atliq Grands	Atliq Palace	Atliq Seasons
W 19	1745	1771	1897	1907	1411	1913	321
W 20	1746	1789	1906	1906	1384	1908	319
W 21	1441	1465	1586	1546	1134	1600	270
W 22	1726	1754	1896	1913	1389	1933	323
W 23	1421	1475	1559	1624	1149	1591	270
W 24	1756	1784	1916	1937	1394	1929	325
W 25	1734	1779	1897	1895	1381	1930	327
W 26	1424	1459	1568	1574	1141	1581	270
W 27	1758	1777	1883	1911	1379	1933	321
W 28	1753	1781	1867	1902	1380	1924	322
W 29	1750	1784	1922	1921	1408	1909	324
W 30	1427	1457	1569	1569	1145	1585	268
W 31	1451	1450	1566	1557	1130	1593	274



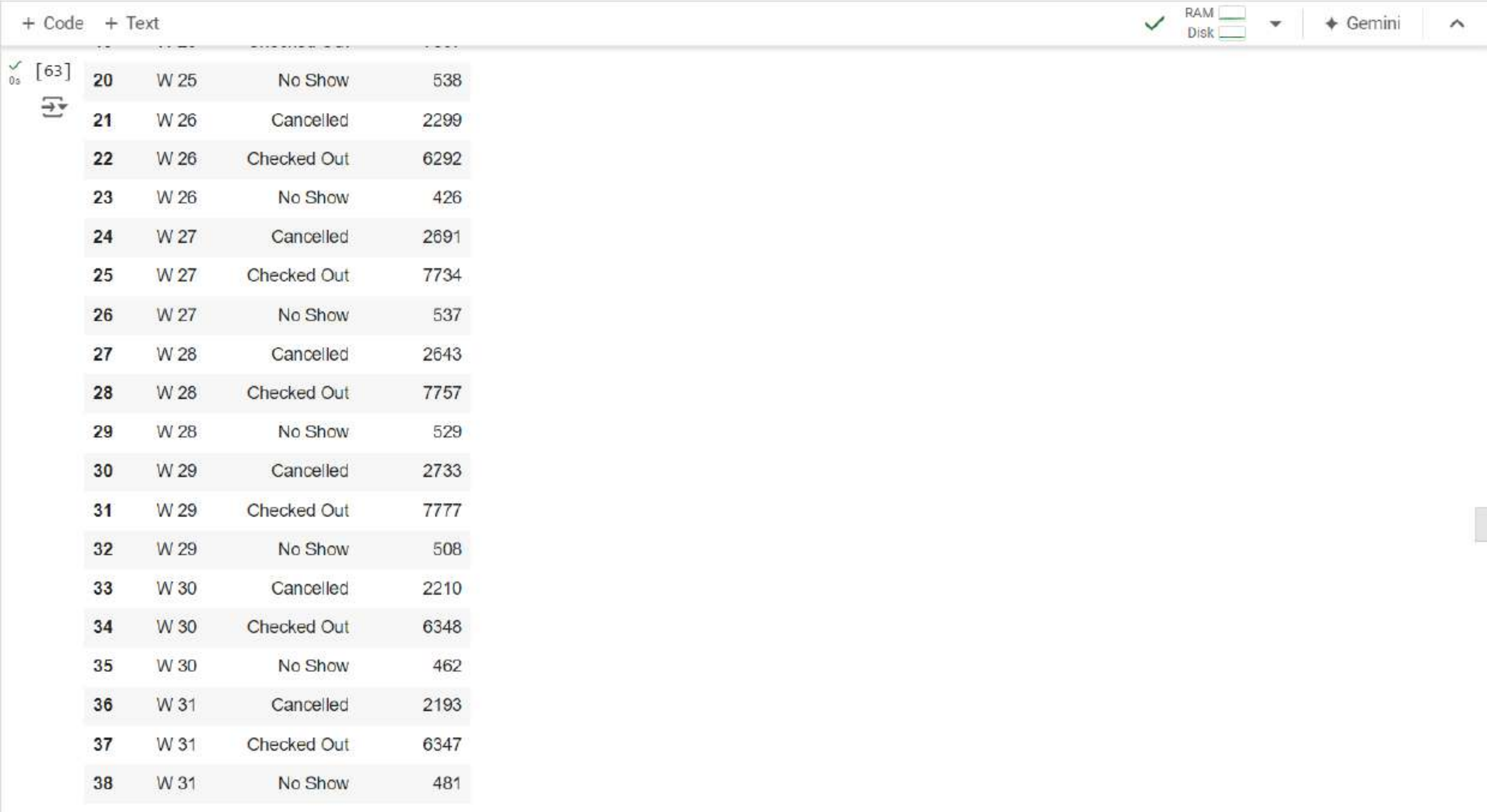
```
weekly_bookings_pivot.plot(kind='barh',legend='false')
plt.ylabel('Total Bookings')
plt.title('Weekly Hotel Bookings')
```

```
Text(0.5, 1.0, 'Weekly Hotel Bookings')
```

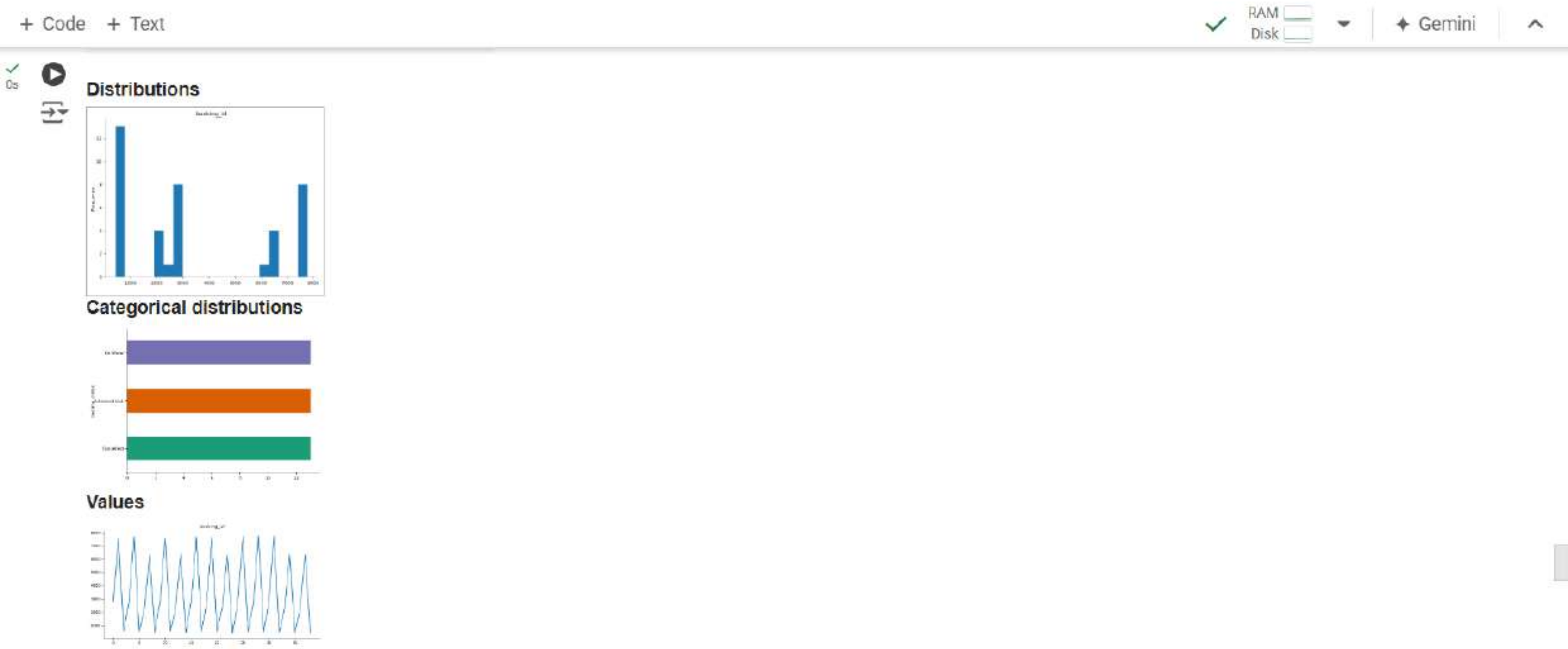
```
x=revenue_tr.groupby(['week no','booking_status']).agg({'booking_id':'nunique'})
x.reset_index(inplace=True)
x
```



	week no	booking_status	booking_id
0	W 19	Cancelled	2779
1	W 19	Checked Out	7628
2	W 19	No Show	558
3	W 20	Cancelled	2711
4	W 20	Checked Out	7699
5	W 20	No Show	548
6	W 21	Cancelled	2226
7	W 21	Checked Out	6331
8	W 21	No Show	485
9	W 22	Cancelled	2772
10	W 22	Checked Out	7585
11	W 22	No Show	577
12	W 23	Cancelled	2222
13	W 23	Checked Out	6399
14	W 23	No Show	468
15	W 24	Cancelled	2797
16	W 24	Checked Out	7688
17	W 24	No Show	556
18	W 25	Cancelled	2748
19	W 25	Checked Out	7657



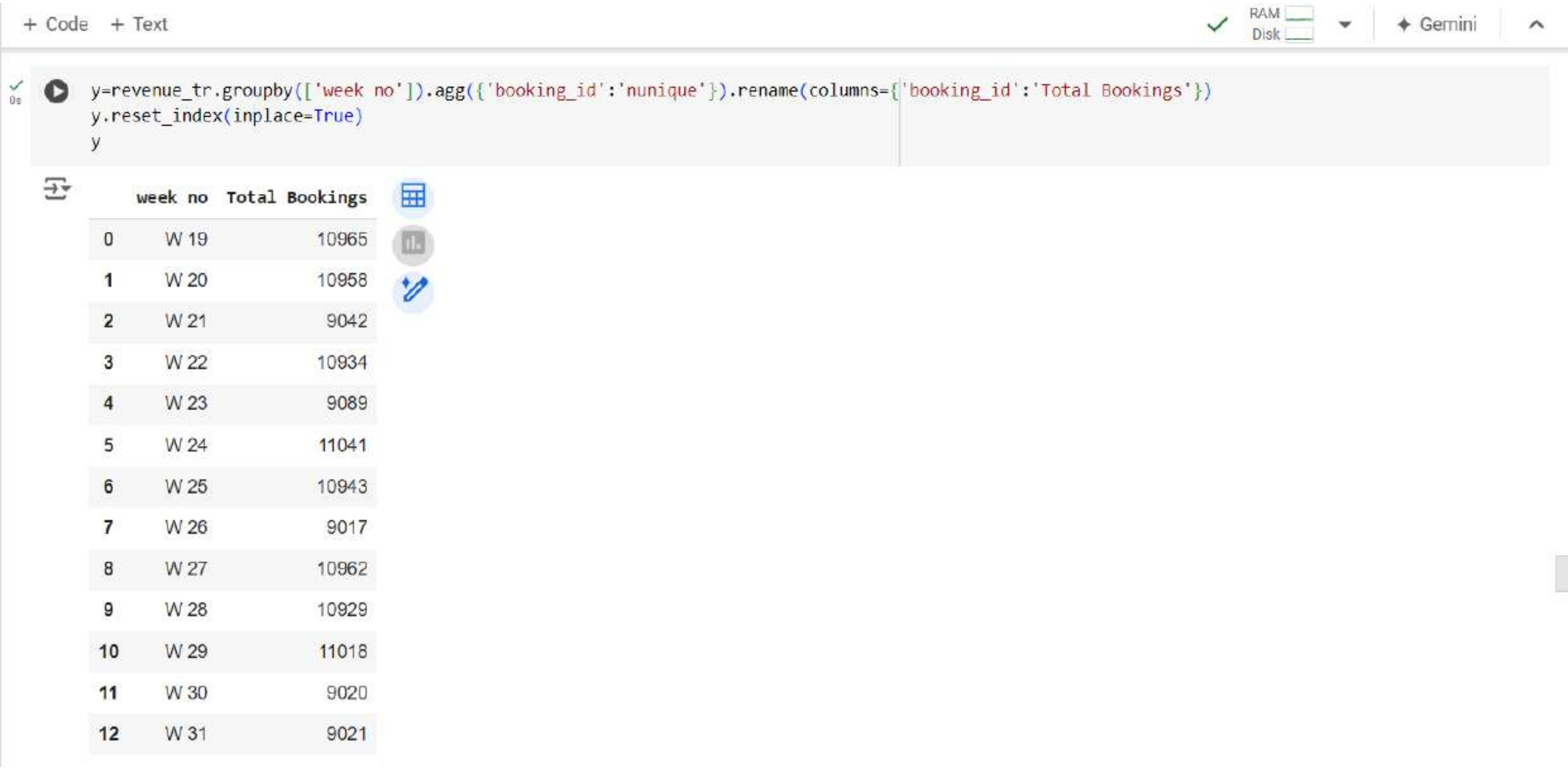
20	W 25	No Show	538
21	W 26	Cancelled	2299
22	W 26	Checked Out	6292
23	W 26	No Show	426
24	W 27	Cancelled	2691
25	W 27	Checked Out	7734
26	W 27	No Show	537
27	W 28	Cancelled	2643
28	W 28	Checked Out	7757
29	W 28	No Show	529
30	W 29	Cancelled	2733
31	W 29	Checked Out	7777
32	W 29	No Show	508
33	W 30	Cancelled	2210
34	W 30	Checked Out	6348
35	W 30	No Show	462
36	W 31	Cancelled	2193
37	W 31	Checked Out	6347
38	W 31	No Show	481



Distributions

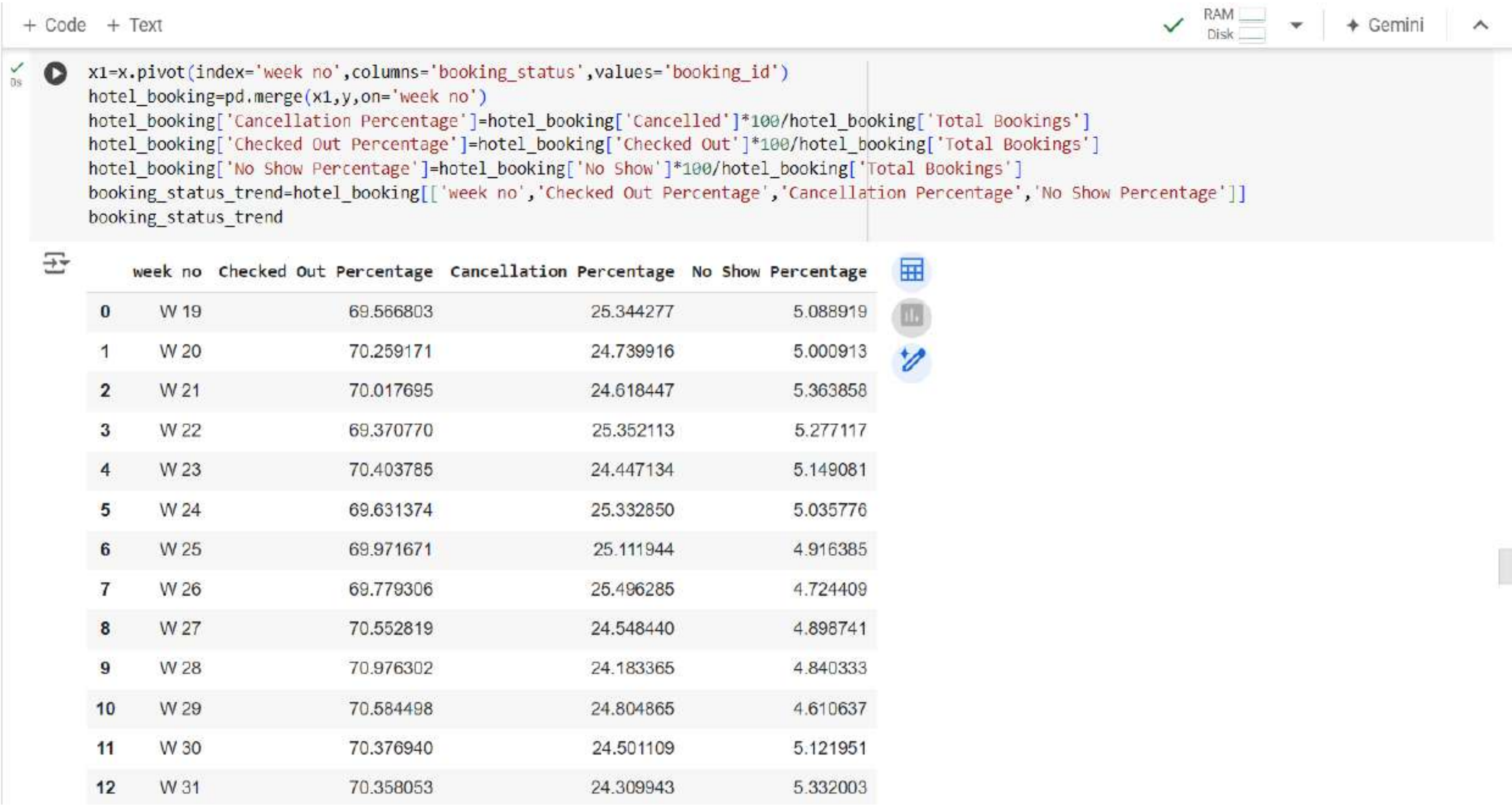
Categorical distributions

Values



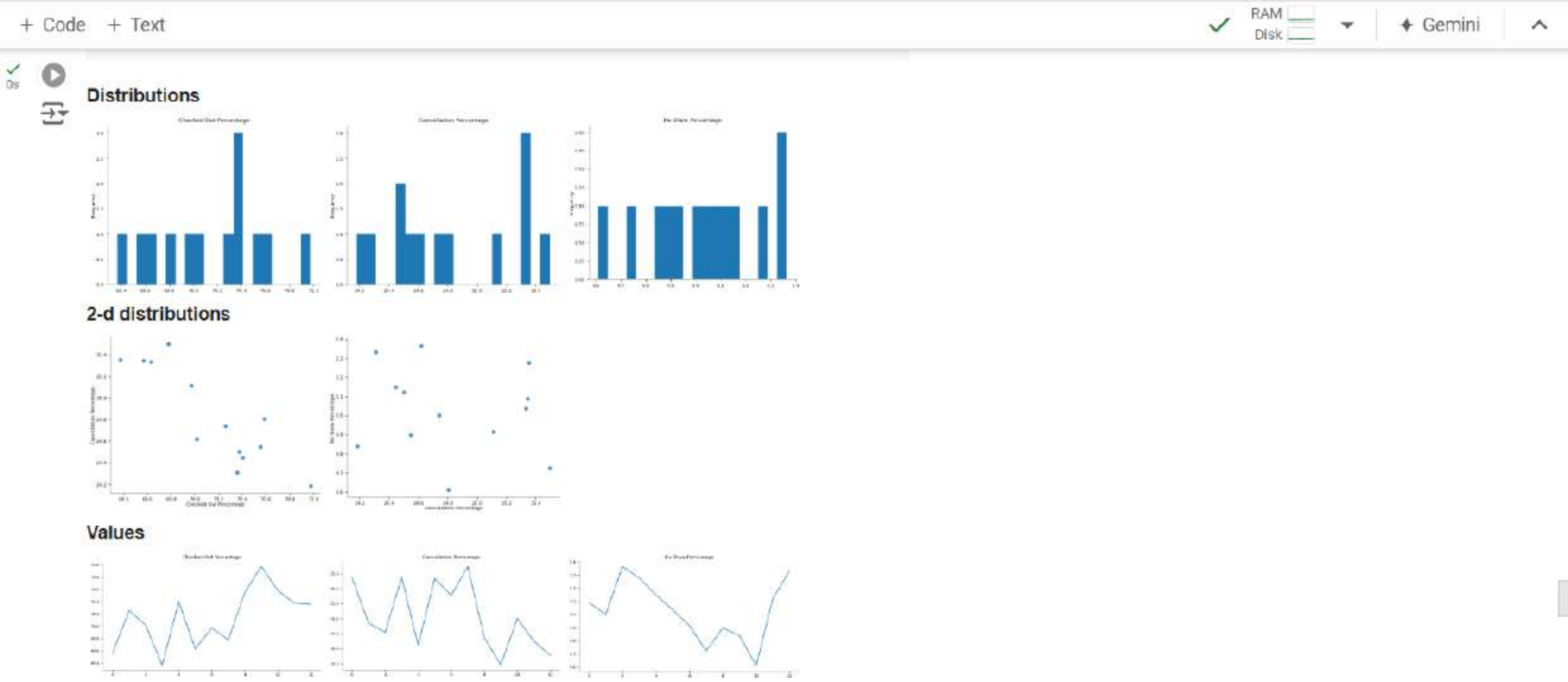
```
y=revenue_tr.groupby(['week no']).agg({'booking_id':'nunique'}).rename(columns={'booking_id':'Total Bookings'})
y.reset_index(inplace=True)
y
```

	week no	Total Bookings
0	W 19	10965
1	W 20	10958
2	W 21	9042
3	W 22	10934
4	W 23	9089
5	W 24	11041
6	W 25	10943
7	W 26	9017
8	W 27	10962
9	W 28	10929
10	W 29	11018
11	W 30	9020
12	W 31	9021



```
x1=x.pivot(index='week no',columns='booking_status',values='booking_id')
hotel_booking=pd.merge(x1,y,on='week no')
hotel_booking['Cancellation Percentage']=hotel_booking['Cancelled']*100/hotel_booking['Total Bookings']
hotel_booking['Checked Out Percentage']=hotel_booking['Checked Out']*100/hotel_booking['Total Bookings']
hotel_booking['No Show Percentage']=hotel_booking['No Show']*100/hotel_booking['Total Bookings']
booking_status_trend=hotel_booking[['week no','Checked Out Percentage','Cancellation Percentage','No Show Percentage']]
booking_status_trend
```

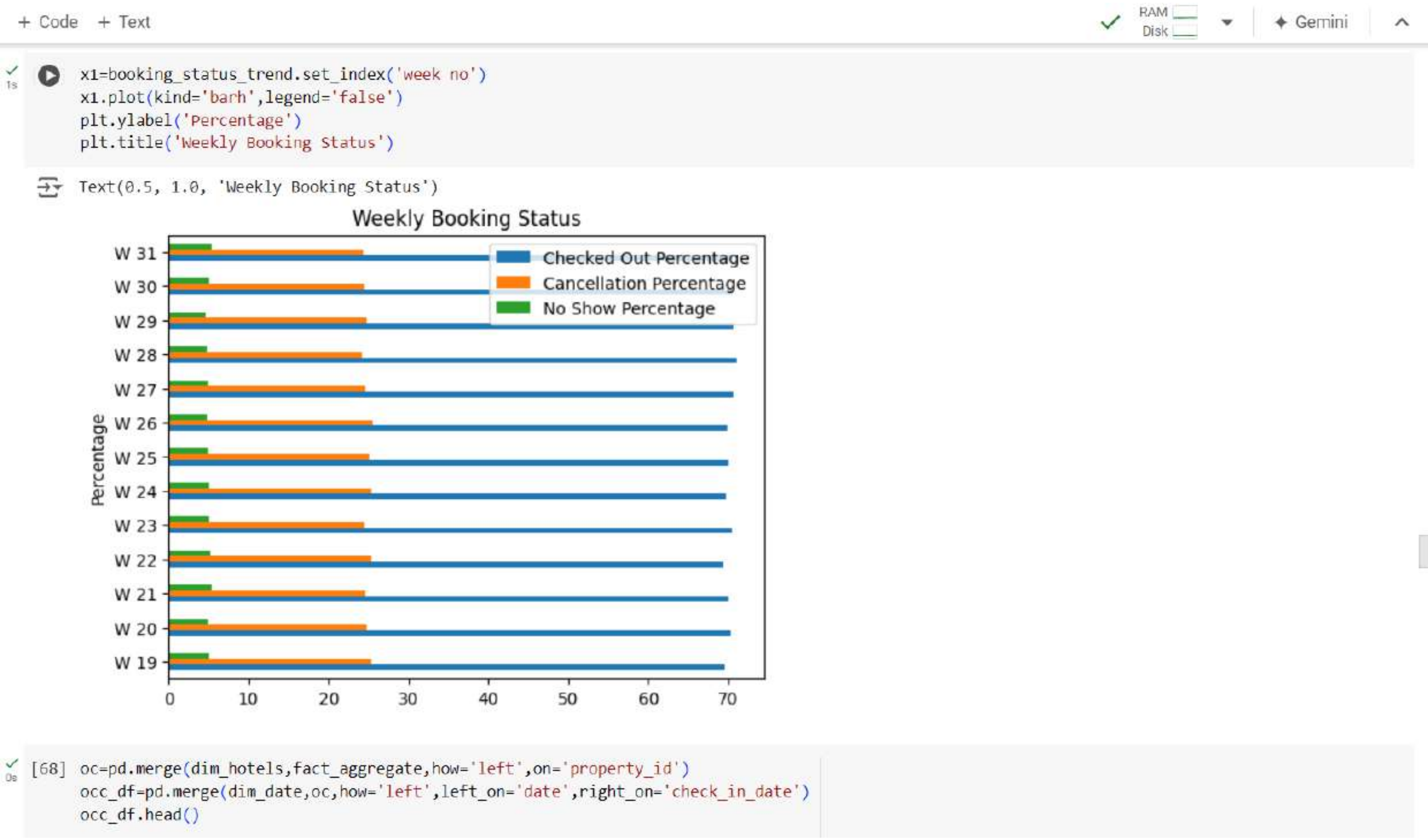
	week no	Checked Out Percentage	Cancellation Percentage	No Show Percentage
0	W 19	69.566803	25.344277	5.088919
1	W 20	70.259171	24.739916	5.000913
2	W 21	70.017695	24.618447	5.363858
3	W 22	69.370770	25.352113	5.277117
4	W 23	70.403785	24.447134	5.149081
5	W 24	69.631374	25.332850	5.035776
6	W 25	69.971671	25.111944	4.916385
7	W 26	69.779306	25.496285	4.724409
8	W 27	70.552819	24.548440	4.898741
9	W 28	70.976302	24.183365	4.840333
10	W 29	70.584498	24.804865	4.610637
11	W 30	70.376940	24.501109	5.121951
12	W 31	70.358053	24.309943	5.332003



Distributions

2-d distributions

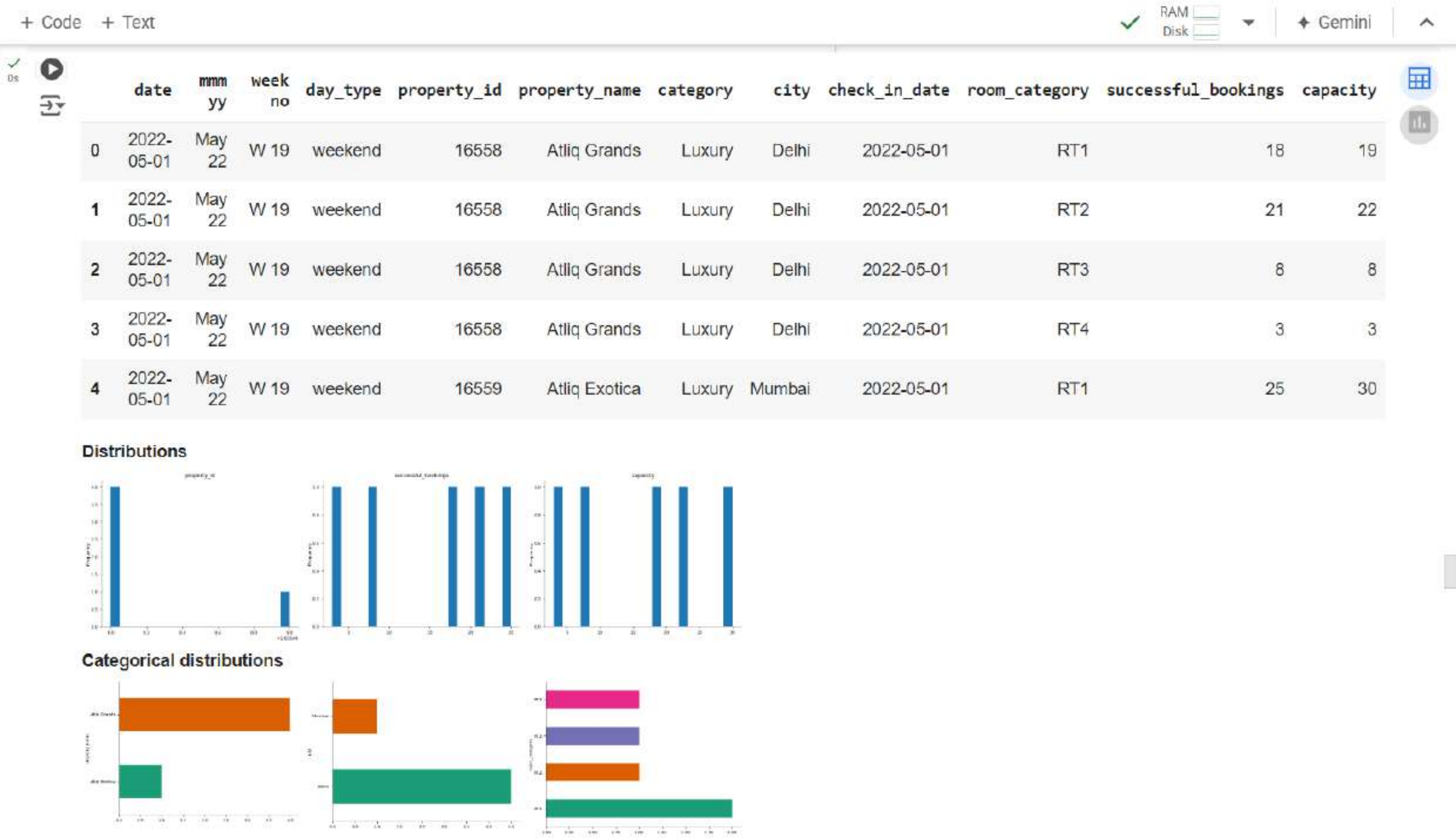
Values



```
x1=booking_status_trend.set_index('week no')
x1.plot(kind='barh',legend='false')
plt.ylabel('Percentage')
plt.title('Weekly Booking Status')
```

```
Text(0.5, 1.0, 'Weekly Booking Status')
```

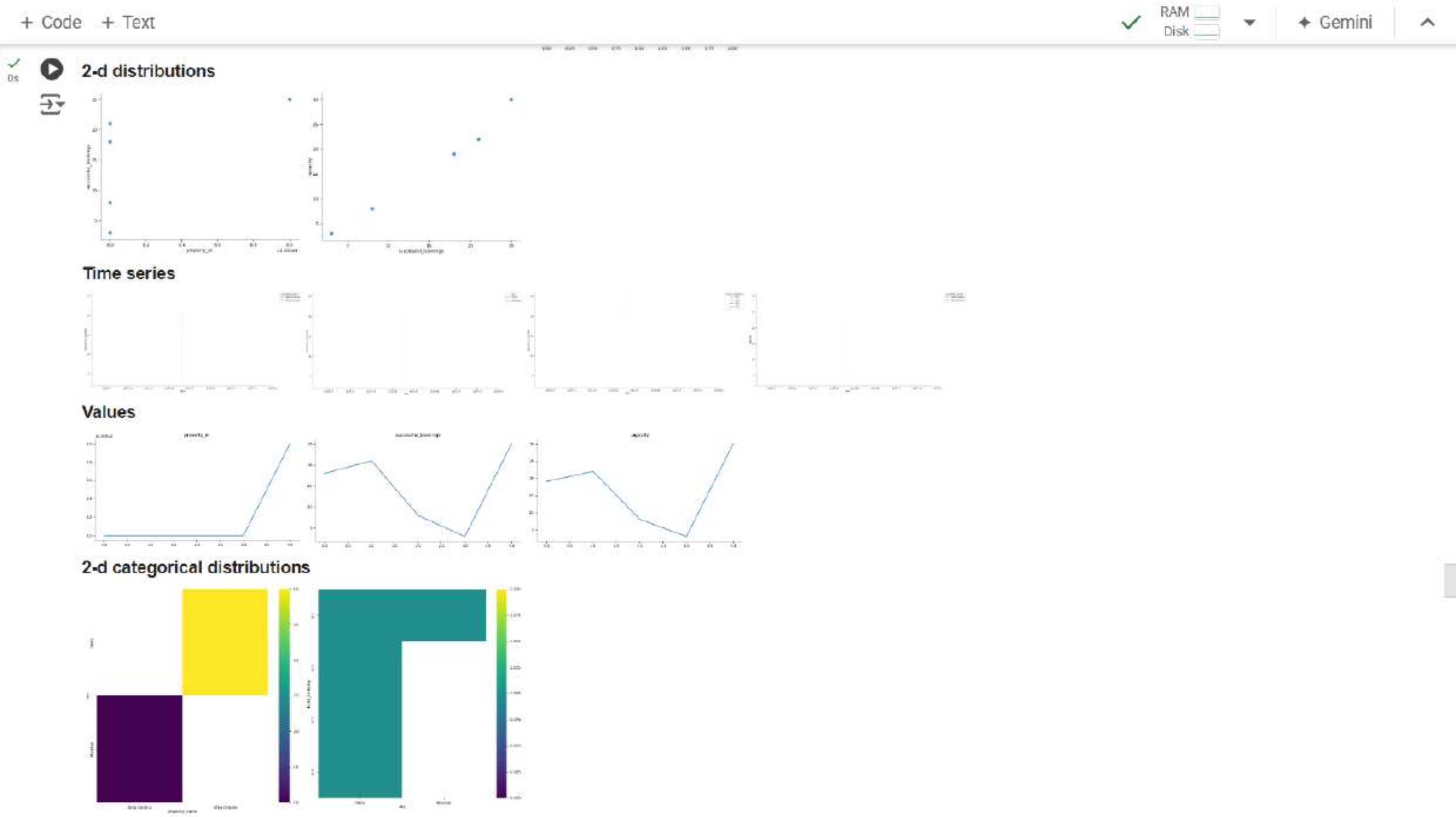
```
[68] oc=pd.merge(dim_hotels,fact_aggregate,how='left',on='property_id')
     occ_df=pd.merge(dim_date,oc,how='left',left_on='date',right_on='check_in_date')
     occ_df.head()
```



	date	mmm yy	week no	day_type	property_id	property_name	category	city	check_in_date	room_category	successful_bookings	capacity
0	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022-05-01	RT1	18	19
1	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022-05-01	RT2	21	22
2	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022-05-01	RT3	8	8
3	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022-05-01	RT4	3	3
4	2022-05-01	May 22	W 19	weekend	16559	Atliq Exotica	Luxury	Mumbai	2022-05-01	RT1	25	30

Distributions

Categorical distributions

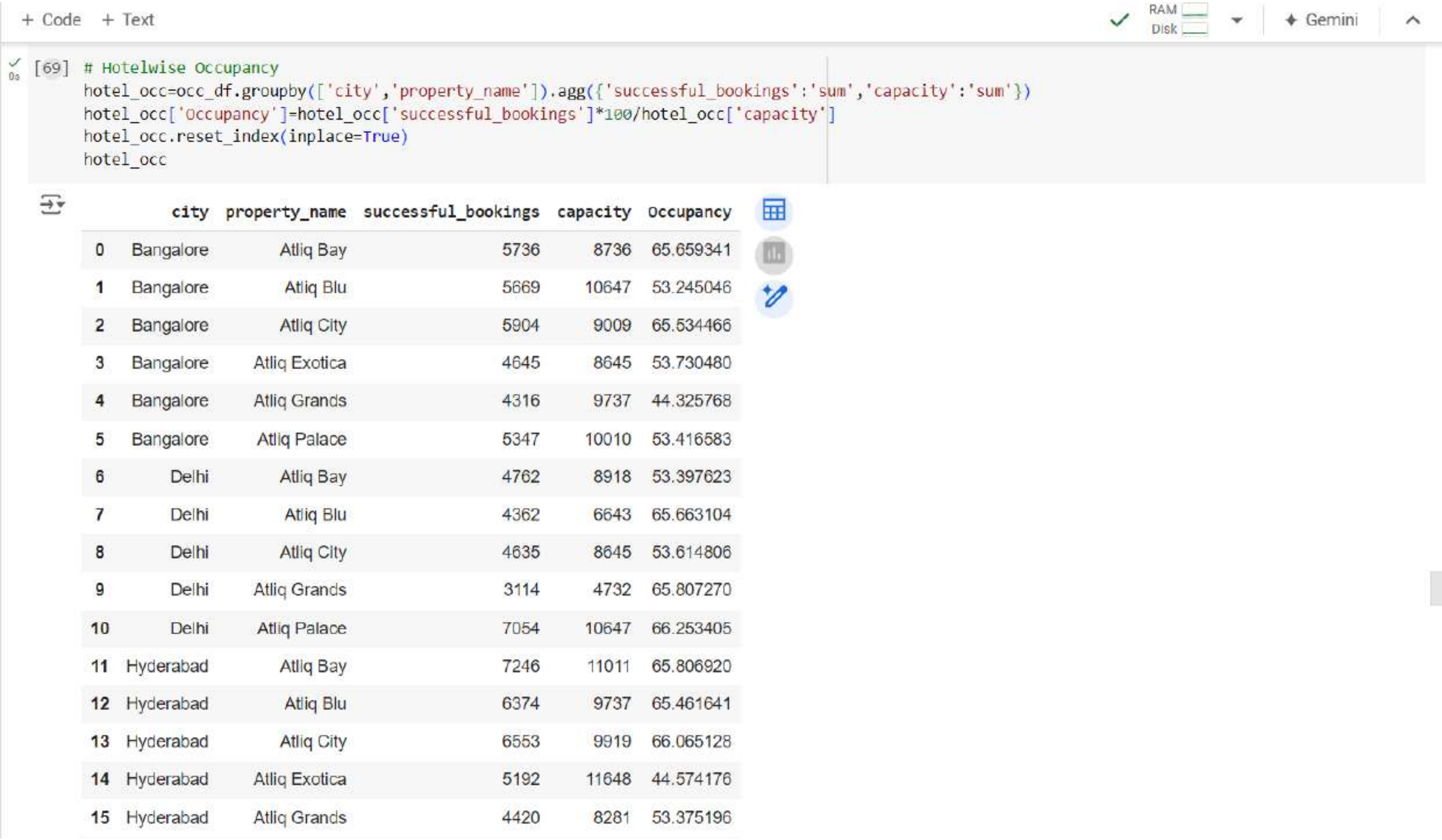


2-d distributions

Time series

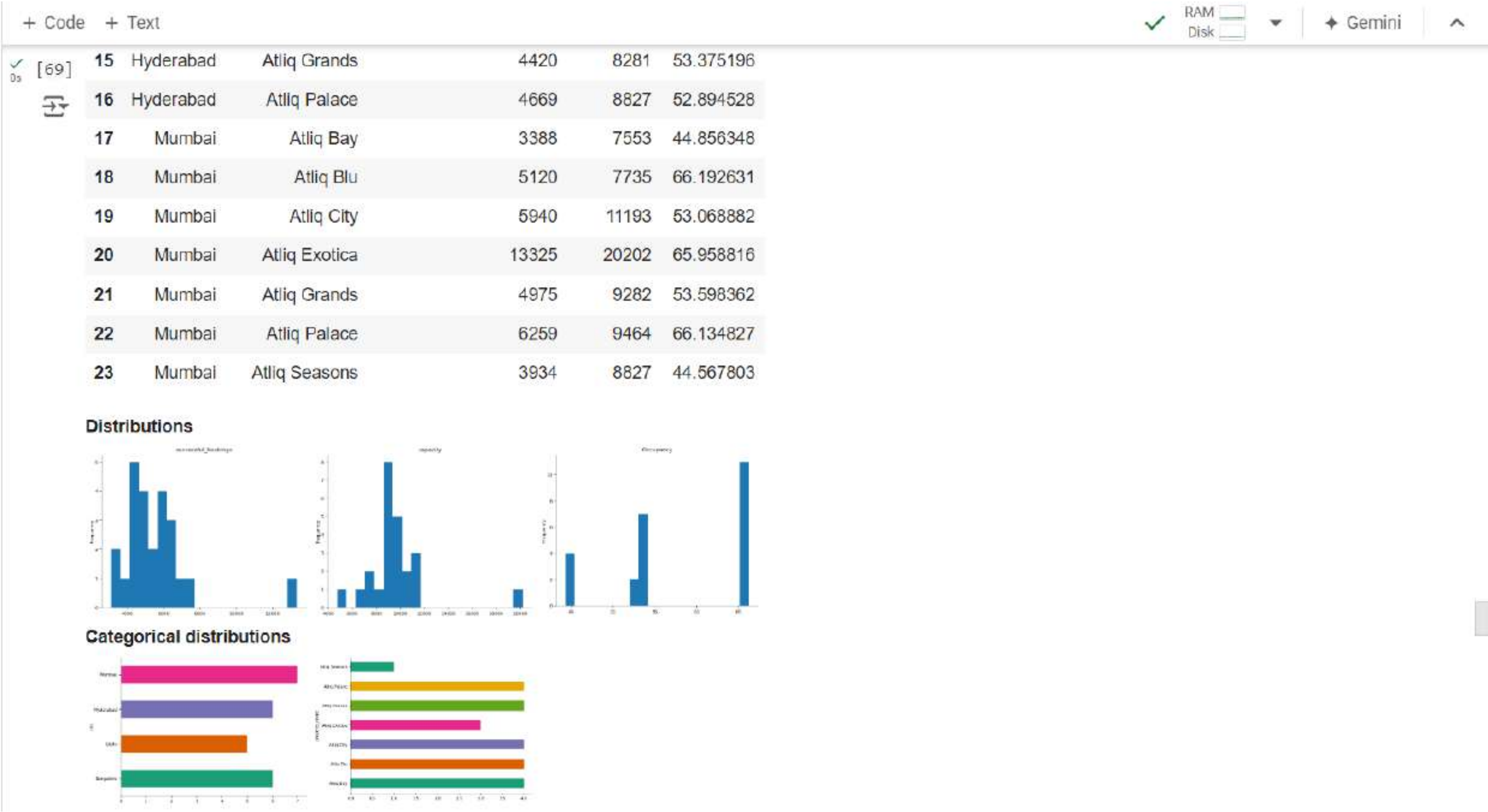
Values

2-d categorical distributions



```
[69] # Hotelwise Occupancy
     hotel_occ=occ_df.groupby(['city','property_name']).agg({'successful_bookings':'sum','capacity':'sum'})
     hotel_occ['Occupancy']=hotel_occ['successful_bookings']*100/hotel_occ['capacity']
     hotel_occ.reset_index(inplace=True)
     hotel_occ
```

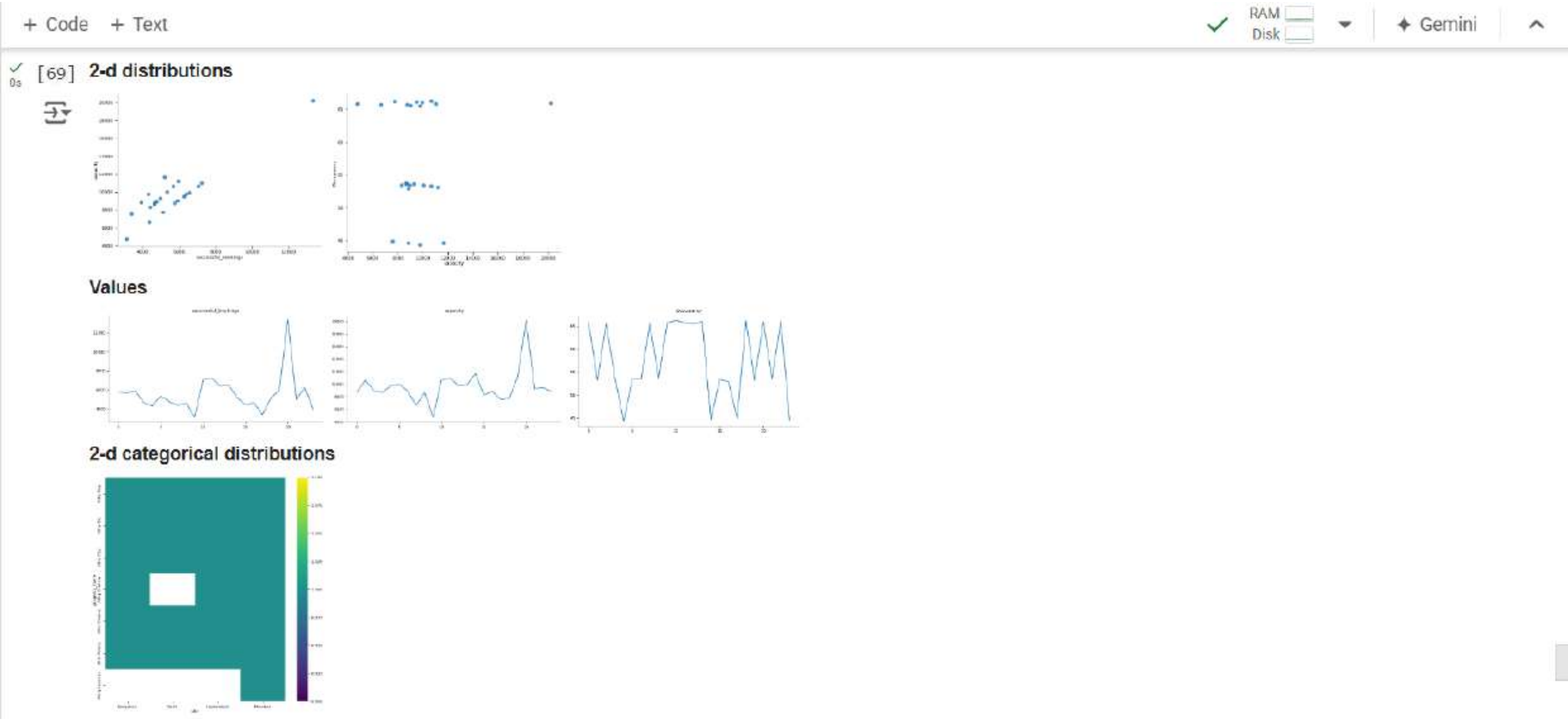
	city	property_name	successful_bookings	capacity	Occupancy
0	Bangalore	Atliq Bay	5736	8736	65.659341
1	Bangalore	Atliq Blu	5669	10647	53.245046
2	Bangalore	Atliq City	5904	9009	65.534466
3	Bangalore	Atliq Exotica	4645	8645	53.730480
4	Bangalore	Atliq Grands	4316	9737	44.325768
5	Bangalore	Atliq Palace	5347	10010	53.416583
6	Delhi	Atliq Bay	4762	8918	53.397623
7	Delhi	Atliq Blu	4362	6643	65.663104
8	Delhi	Atliq City	4635	8645	53.614806
9	Delhi	Atliq Grands	3114	4732	65.807270
10	Delhi	Atliq Palace	7054	10647	66.253405
11	Hyderabad	Atliq Bay	7246	11011	65.806920
12	Hyderabad	Atliq Blu	6374	9737	65.461641
13	Hyderabad	Atliq City	6553	9919	66.065128
14	Hyderabad	Atliq Exotica	5192	11648	44.574176
15	Hyderabad	Atliq Grands	4420	8281	53.375196



15	Hyderabad	Atliq Grands	4420	8281	53.375196
16	Hyderabad	Atliq Palace	4669	8827	52.894528
17	Mumbai	Atliq Bay	3388	7553	44.856348
18	Mumbai	Atliq Blu	5120	7735	66.192631
19	Mumbai	Atliq City	5940	11193	53.068882
20	Mumbai	Atliq Exotica	13325	20202	65.958816
21	Mumbai	Atliq Grands	4975	9282	53.598362
22	Mumbai	Atliq Palace	6259	9464	66.134827
23	Mumbai	Atliq Seasons	3934	8827	44.567803

Distributions

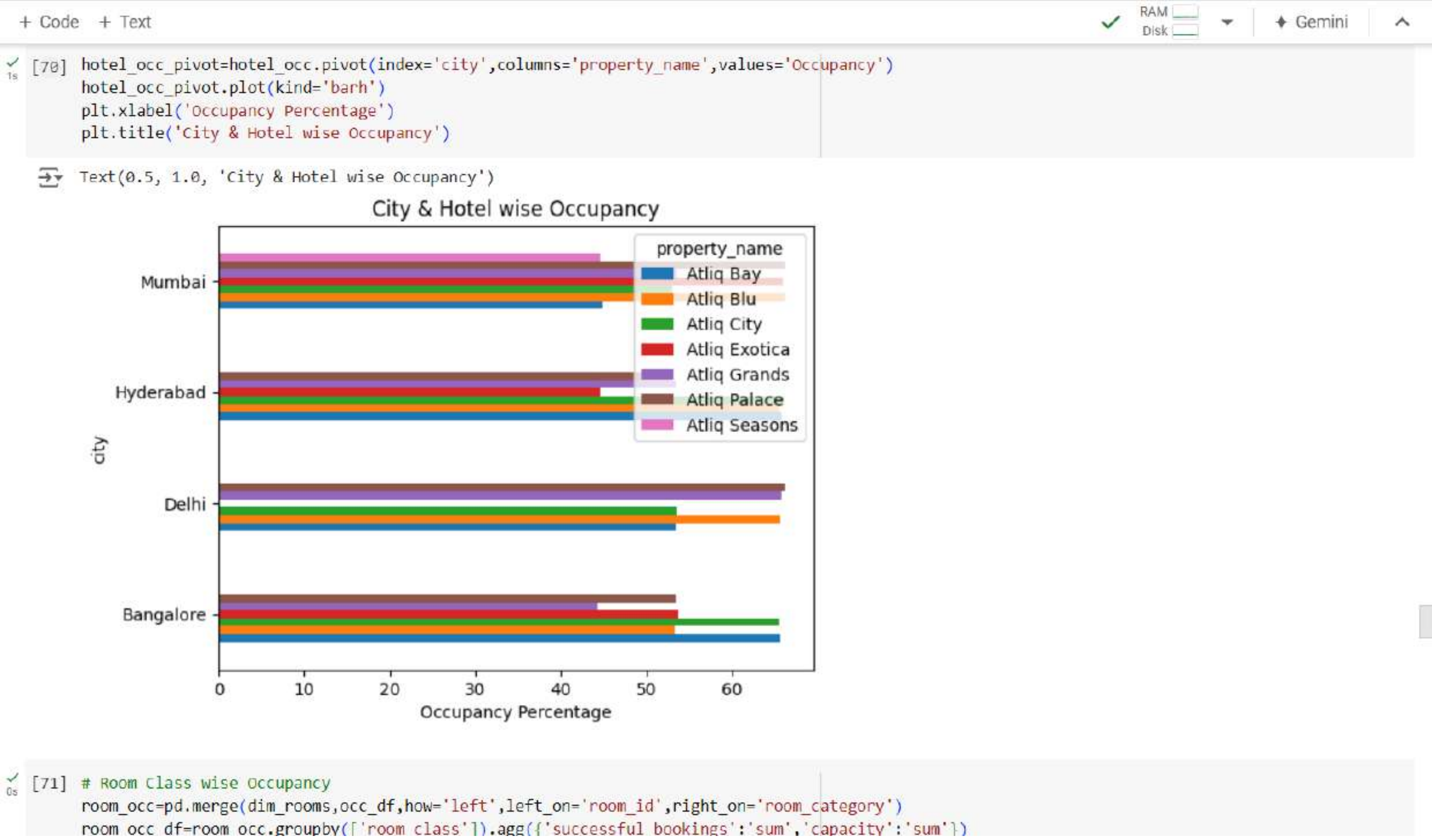
Categorical distributions



2-d distributions

Values

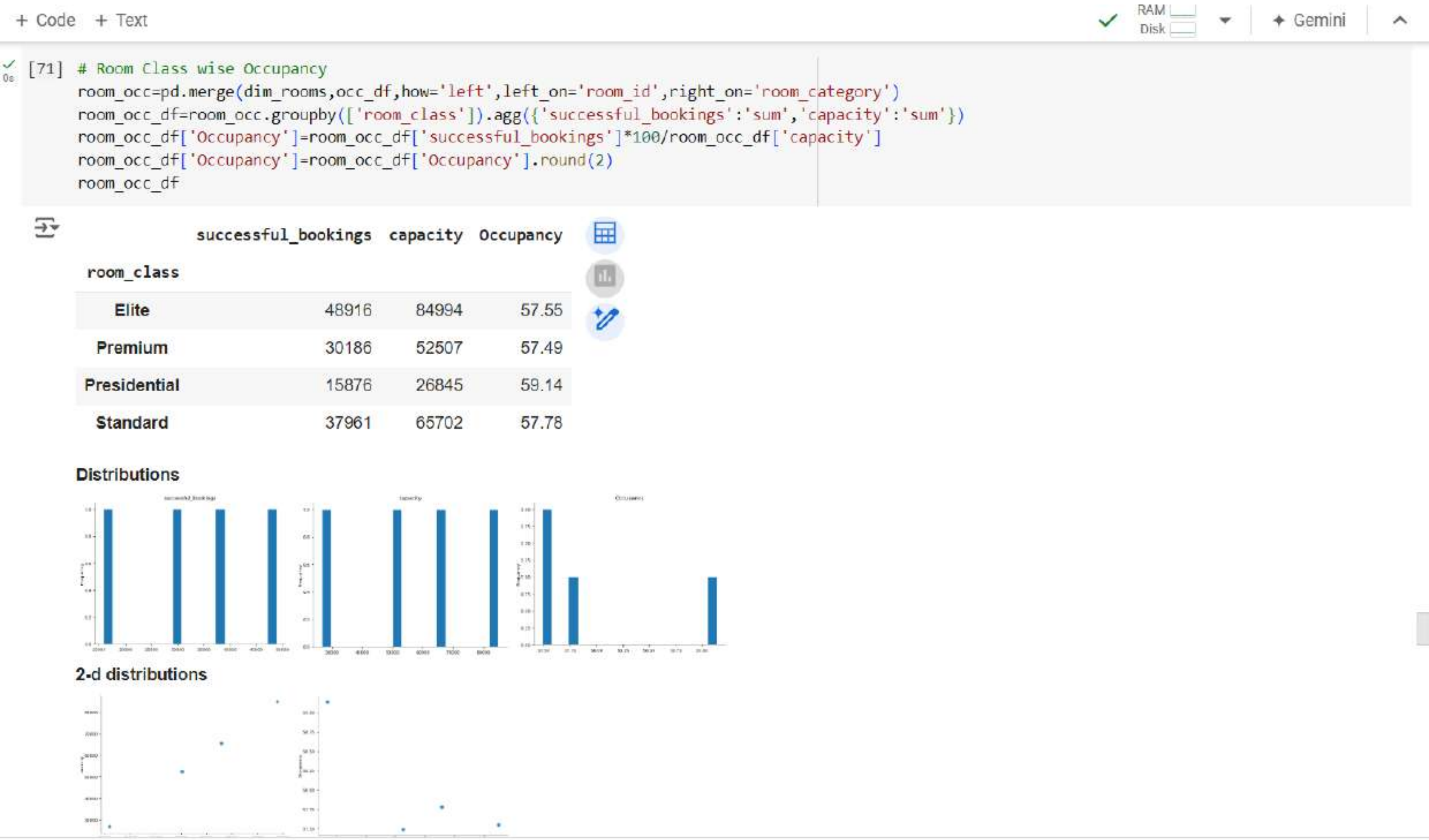
2-d categorical distributions



```
[70] hotel_occ_pivot=hotel_occ.pivot(index='city',columns='property_name',values='Occupancy')
     hotel_occ_pivot.plot(kind='barh')
     plt.xlabel('Occupancy Percentage')
     plt.title('City & Hotel wise Occupancy')
```

```
Text(0.5, 1.0, 'City & Hotel wise Occupancy')
```

```
[71] # Room Class wise Occupancy
     room_occ=pd.merge(dim_rooms,occ_df,how='left',left_on='room_id',right_on='room_category')
     room_occ_df=room_occ.groupby(['room_class']).agg({'successful_bookings':'sum','capacity':'sum'})
```

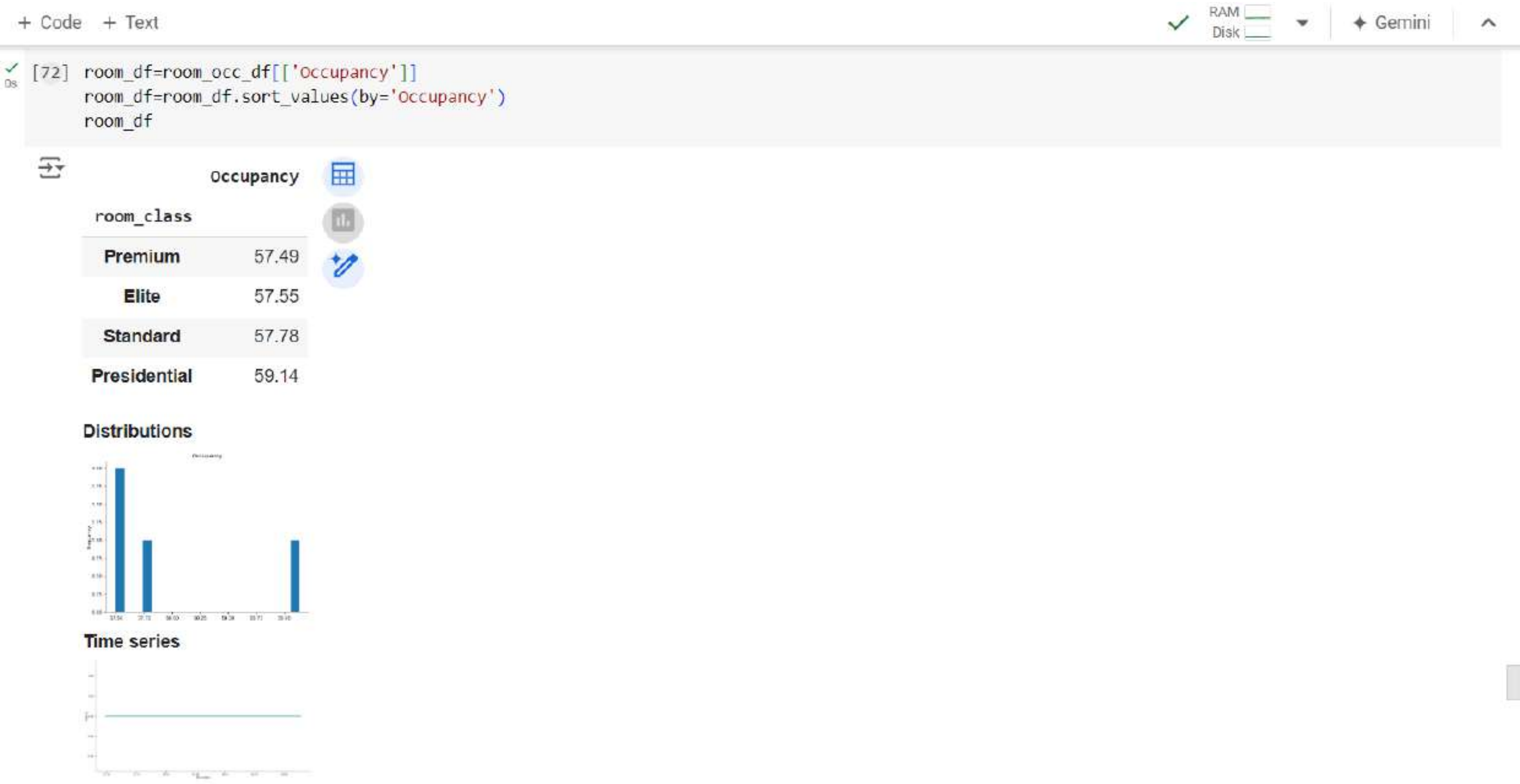


```
# Room Class wise Occupancy
room_occ=pd.merge(dim_rooms,occ_df,how='left',left_on='room_id',right_on='room_category')
room_occ_df=room_occ.groupby(['room_class']).agg({'successful_bookings':'sum','capacity':'sum'})
room_occ_df['Occupancy']=room_occ_df['successful_bookings']*100/room_occ_df['capacity']
room_occ_df['Occupancy']=room_occ_df['Occupancy'].round(2)
room_occ_df
```

	successful_bookings	capacity	Occupancy
room_class			
Elite	48916	84994	57.55
Premium	30186	52507	57.49
Presidential	15876	26845	59.14
Standard	37961	65702	57.78

Distributions

2-d distributions

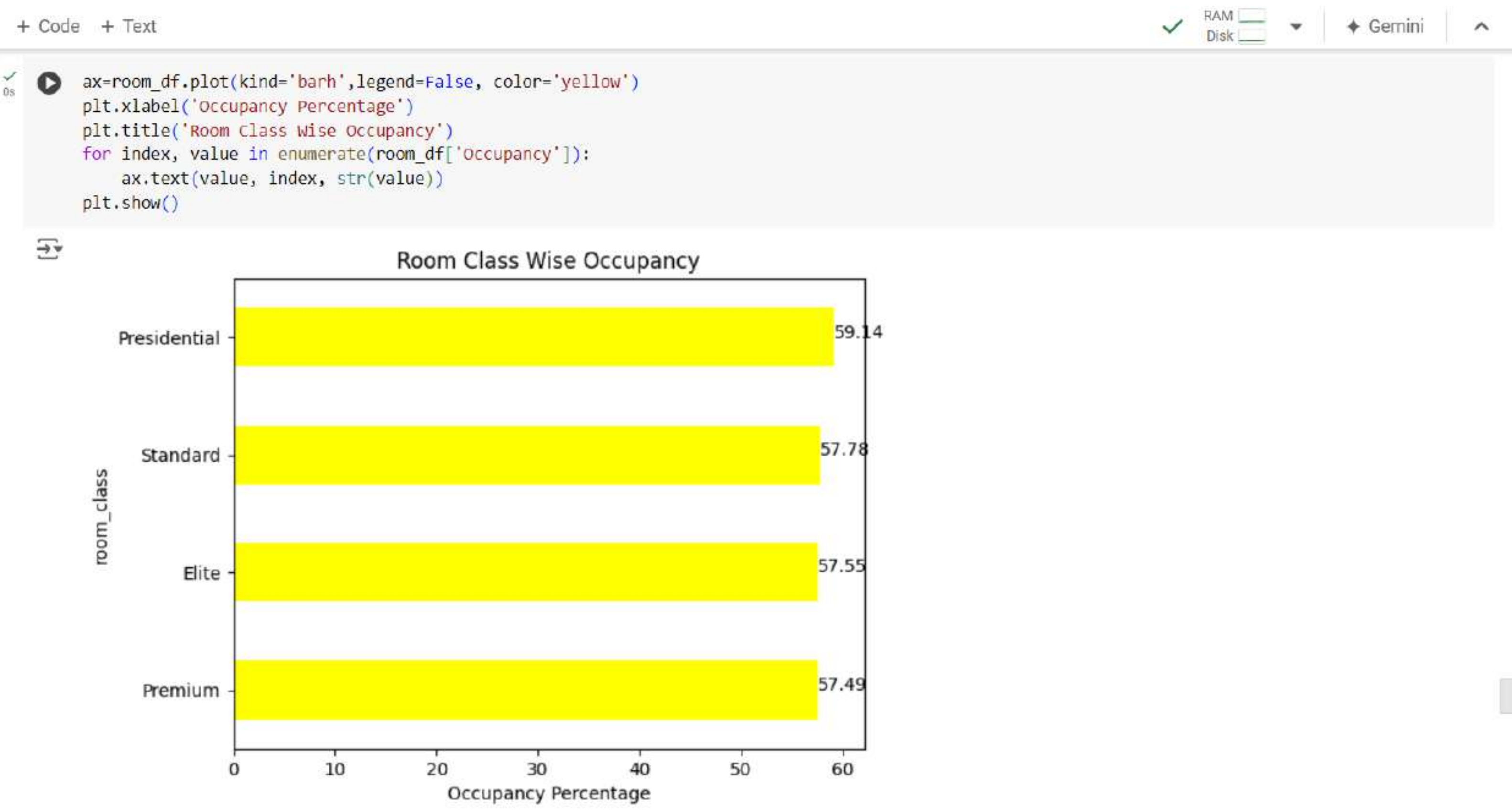


```
room_df=room_occ_df[['Occupancy']]
room_df=room_df.sort_values(by='Occupancy')
room_df
```

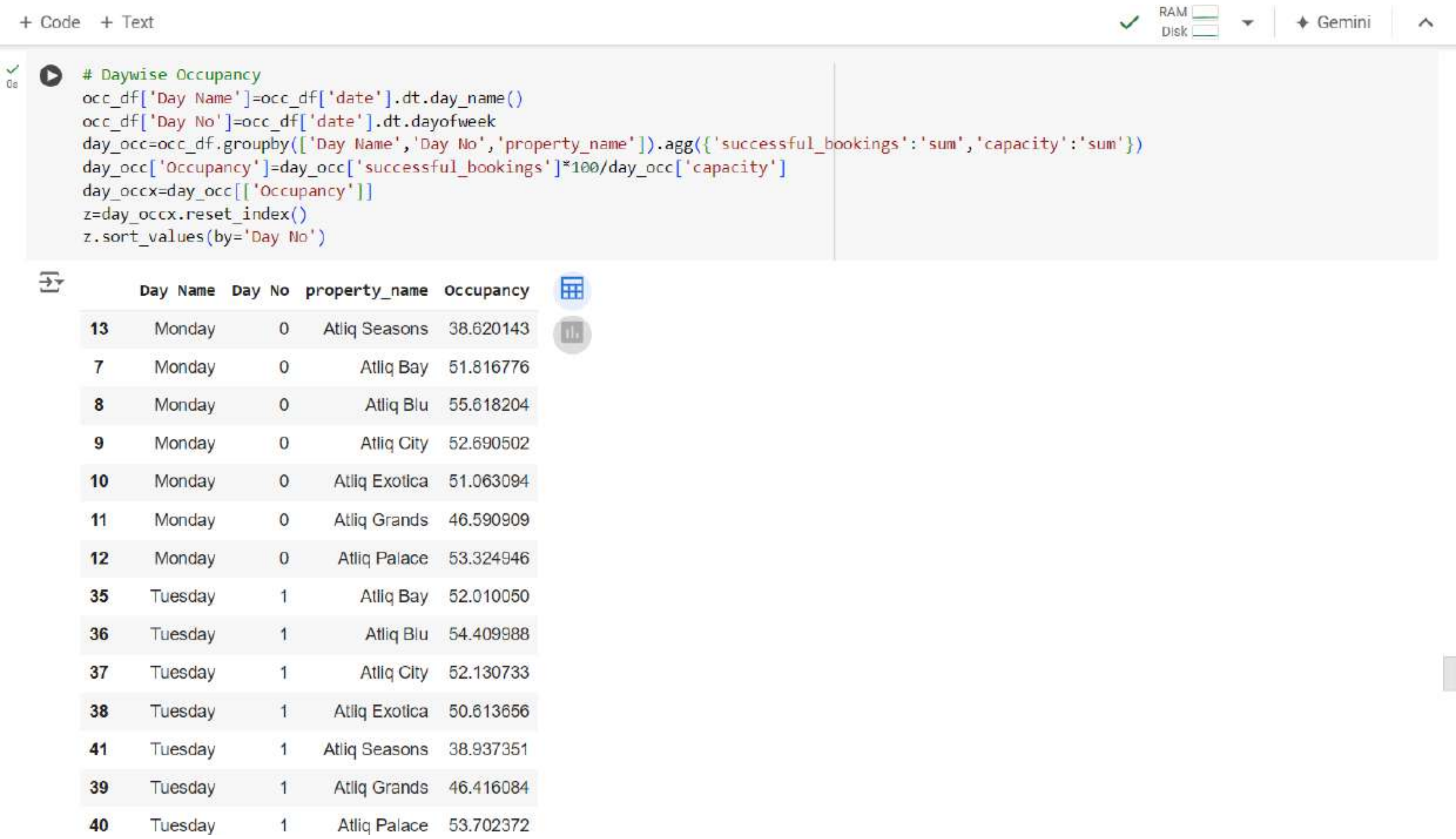
room_class	Occupancy
Premium	57.49
Elite	57.55
Standard	57.78
Presidential	59.14

Distributions

Time series

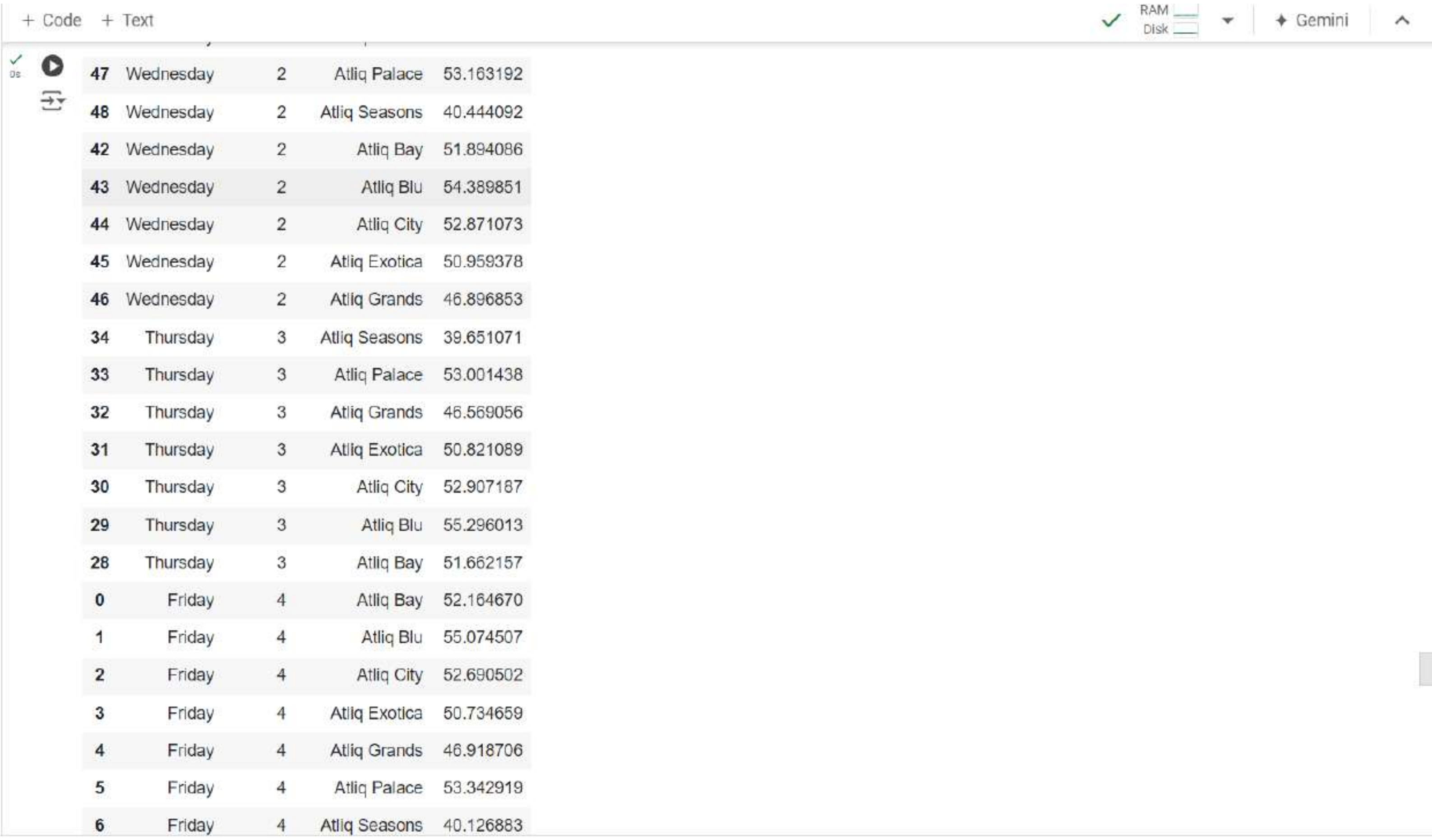


```
ax=room_df.plot(kind='barh',legend=False, color='yellow')
plt.xlabel('Occupancy Percentage')
plt.title('Room Class Wise Occupancy')
for index, value in enumerate(room_df['Occupancy']):
    ax.text(value, index, str(value))
plt.show()
```

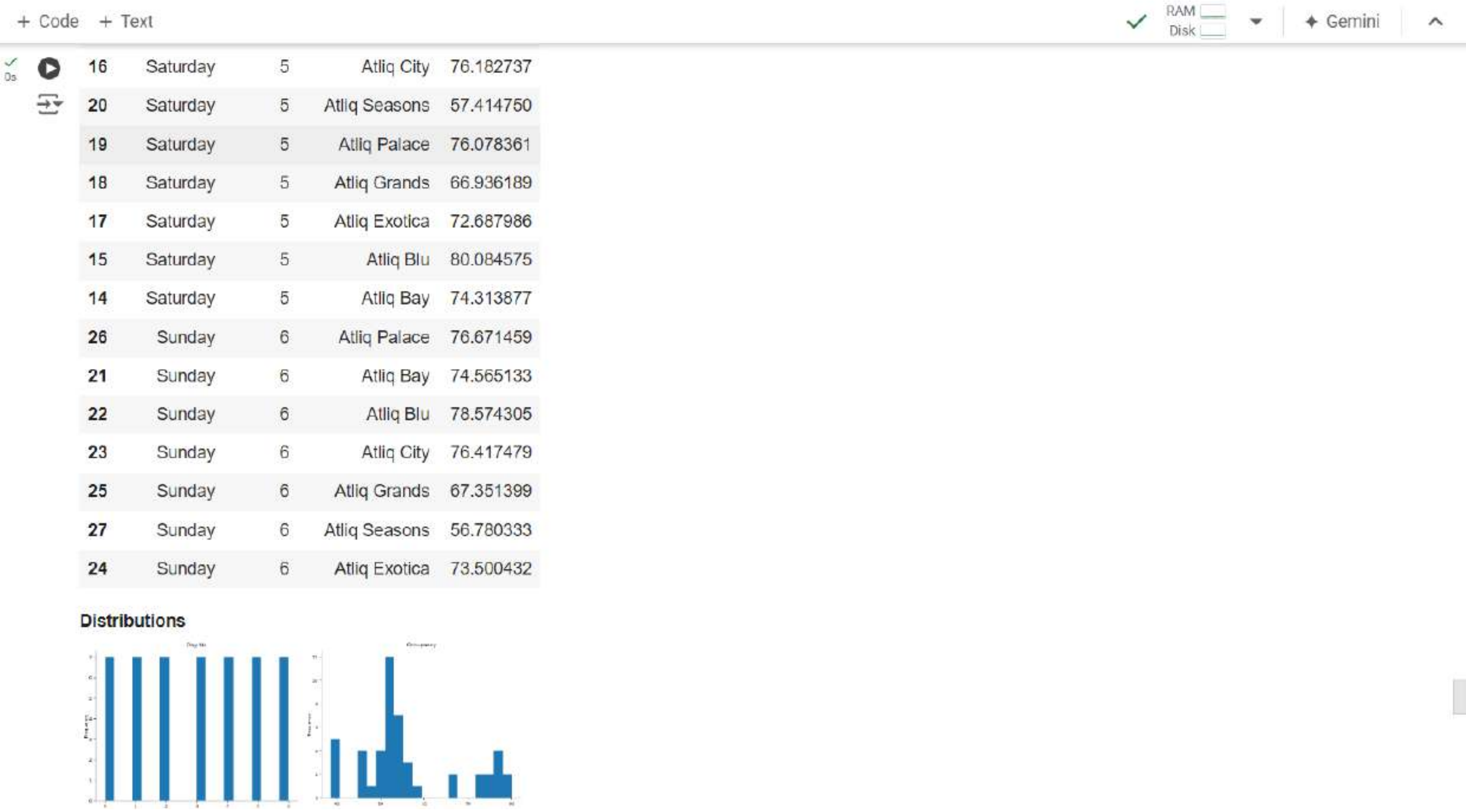


```
# Daywise Occupancy
occ_df['Day Name']=occ_df['date'].dt.day_name()
occ_df['Day No']=occ_df['date'].dt.dayofweek
day_occ=occ_df.groupby(['Day Name','Day No','property_name']).agg({'successful_bookings':'sum','capacity':'sum'})
day_occ['Occupancy']=day_occ['successful_bookings']*100/day_occ['capacity']
day_occx=day_occ[['Occupancy']]
z=day_occx.reset_index()
z.sort_values(by='Day No')
```

```
    Day Name  Day No  property_name  Occupancy
13    Monday       0   Atliq Seasons  38.620143
7     Monday       0       Atliq Bay  51.816776
8     Monday       0       Atliq Blu  55.618204
9     Monday       0      Atliq City  52.690502
10    Monday       0    Atliq Exotica  51.063094
11    Monday       0    Atliq Grands  46.590909
12    Monday       0    Atliq Palace  53.324946
35   Tuesday       1       Atliq Bay  52.010050
36   Tuesday       1       Atliq Blu  54.409988
37   Tuesday       1      Atliq City  52.130733
38   Tuesday       1    Atliq Exotica  50.613656
41   Tuesday       1   Atliq Seasons  38.937351
39   Tuesday       1    Atliq Grands  46.416084
40   Tuesday       1    Atliq Palace  53.702372
```

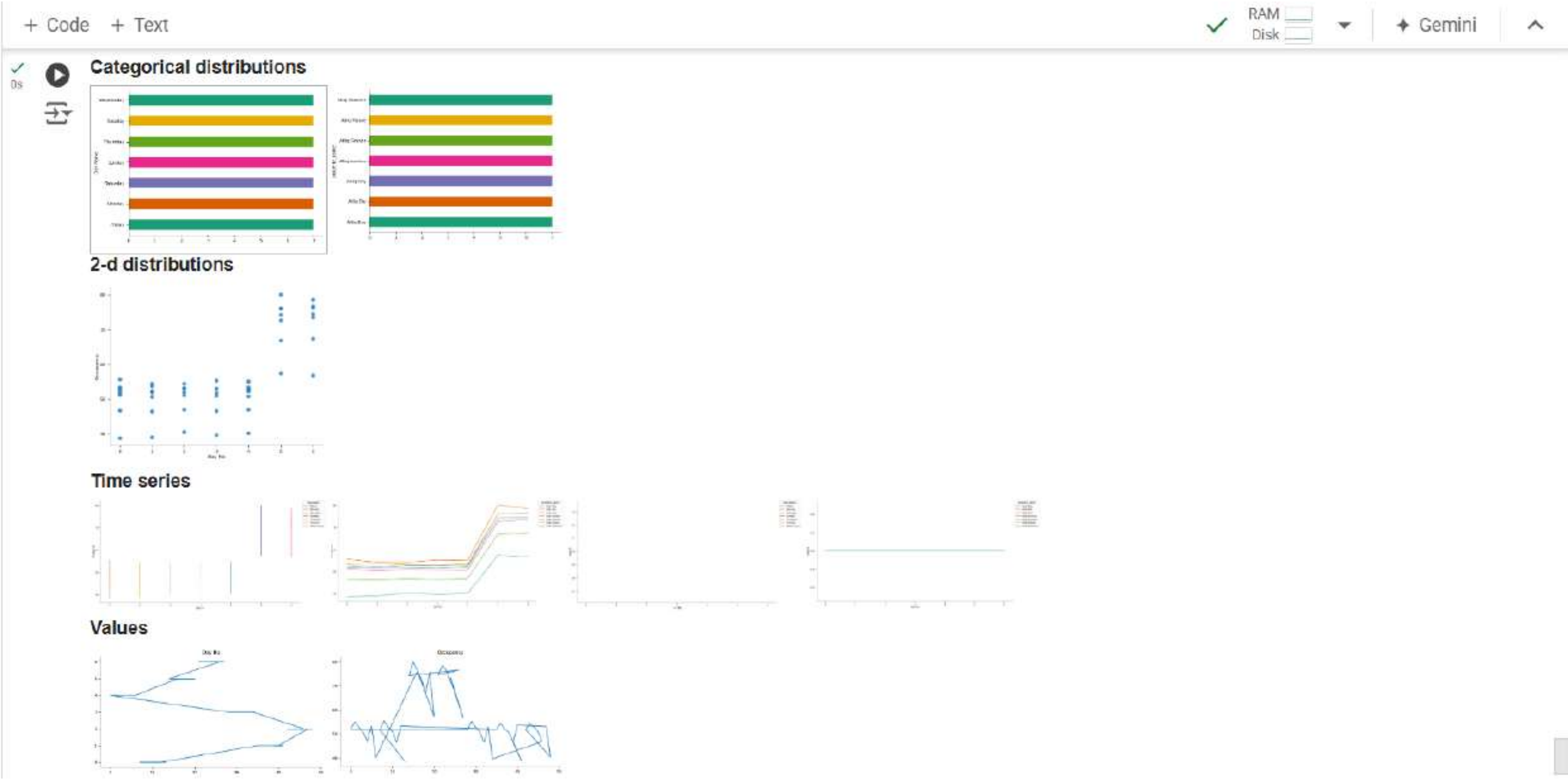


47	Wednesday	2	Atliq Palace	53.163192
48	Wednesday	2	Atliq Seasons	40.444092
42	Wednesday	2	Atliq Bay	51.894086
43	Wednesday	2	Atliq Blu	54.389851
44	Wednesday	2	Atliq City	52.871073
45	Wednesday	2	Atliq Exotica	50.959378
46	Wednesday	2	Atliq Grands	46.896853
34	Thursday	3	Atliq Seasons	39.651071
33	Thursday	3	Atliq Palace	53.001438
32	Thursday	3	Atliq Grands	46.569056
31	Thursday	3	Atliq Exotica	50.821089
30	Thursday	3	Atliq City	52.907187
29	Thursday	3	Atliq Blu	55.296013
28	Thursday	3	Atliq Bay	51.662157
0	Friday	4	Atliq Bay	52.164670
1	Friday	4	Atliq Blu	55.074507
2	Friday	4	Atliq City	52.690502
3	Friday	4	Atliq Exotica	50.734659
4	Friday	4	Atliq Grands	46.918706
5	Friday	4	Atliq Palace	53.342919
6	Friday	4	Atliq Seasons	40.126883



16	Saturday	5	Atliq City	76.182737
20	Saturday	5	Atliq Seasons	57.414750
19	Saturday	5	Atliq Palace	76.078361
18	Saturday	5	Atliq Grands	66.936189
17	Saturday	5	Atliq Exotica	72.687986
15	Saturday	5	Atliq Blu	80.084575
14	Saturday	5	Atliq Bay	74.313877
26	Sunday	6	Atliq Palace	76.671459
21	Sunday	6	Atliq Bay	74.565133
22	Sunday	6	Atliq Blu	78.574305
23	Sunday	6	Atliq City	76.417479
25	Sunday	6	Atliq Grands	67.351399
27	Sunday	6	Atliq Seasons	56.780333
24	Sunday	6	Atliq Exotica	73.500432

Distributions

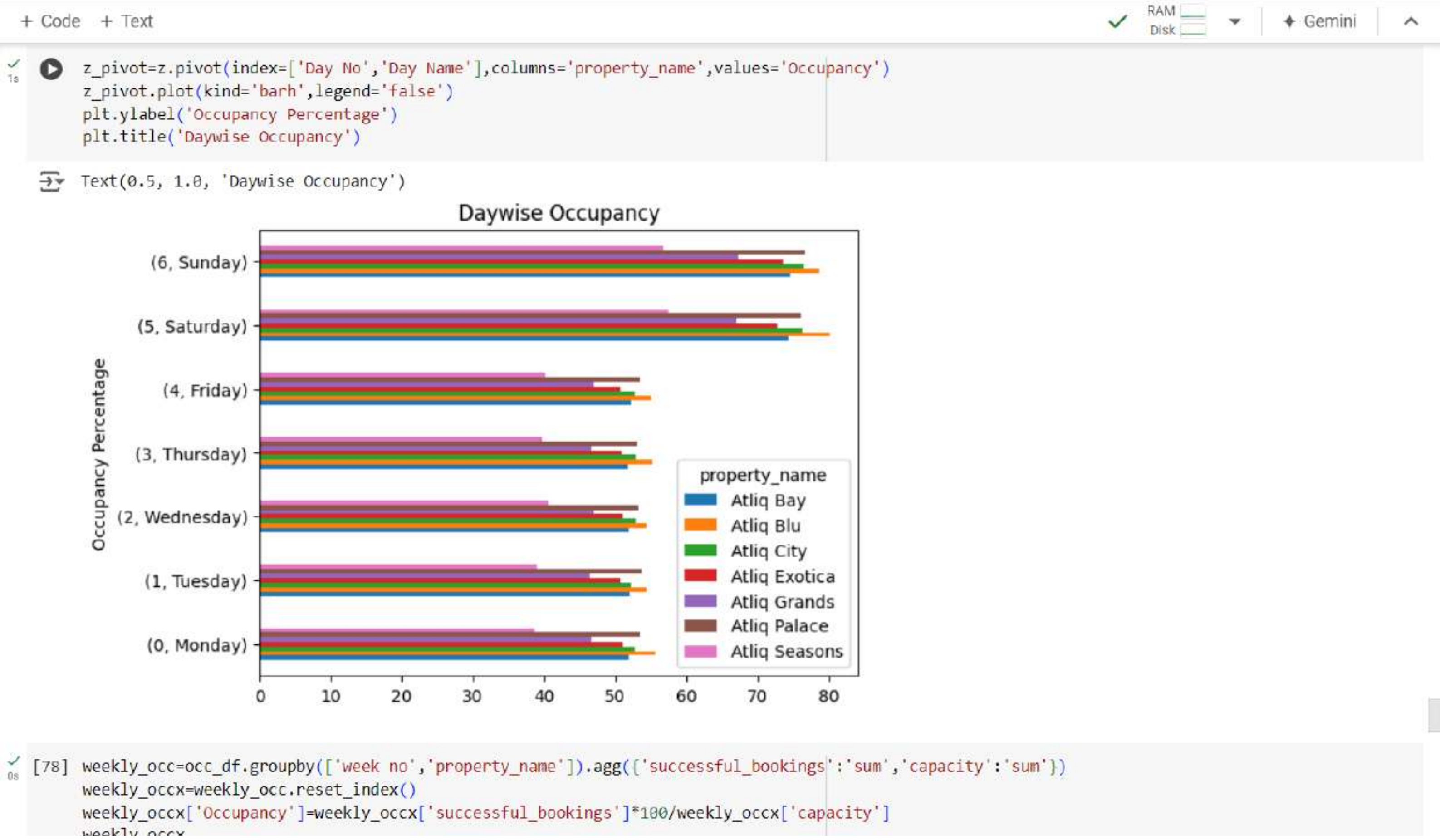


Categorical distributions

2-d distributions

Time series

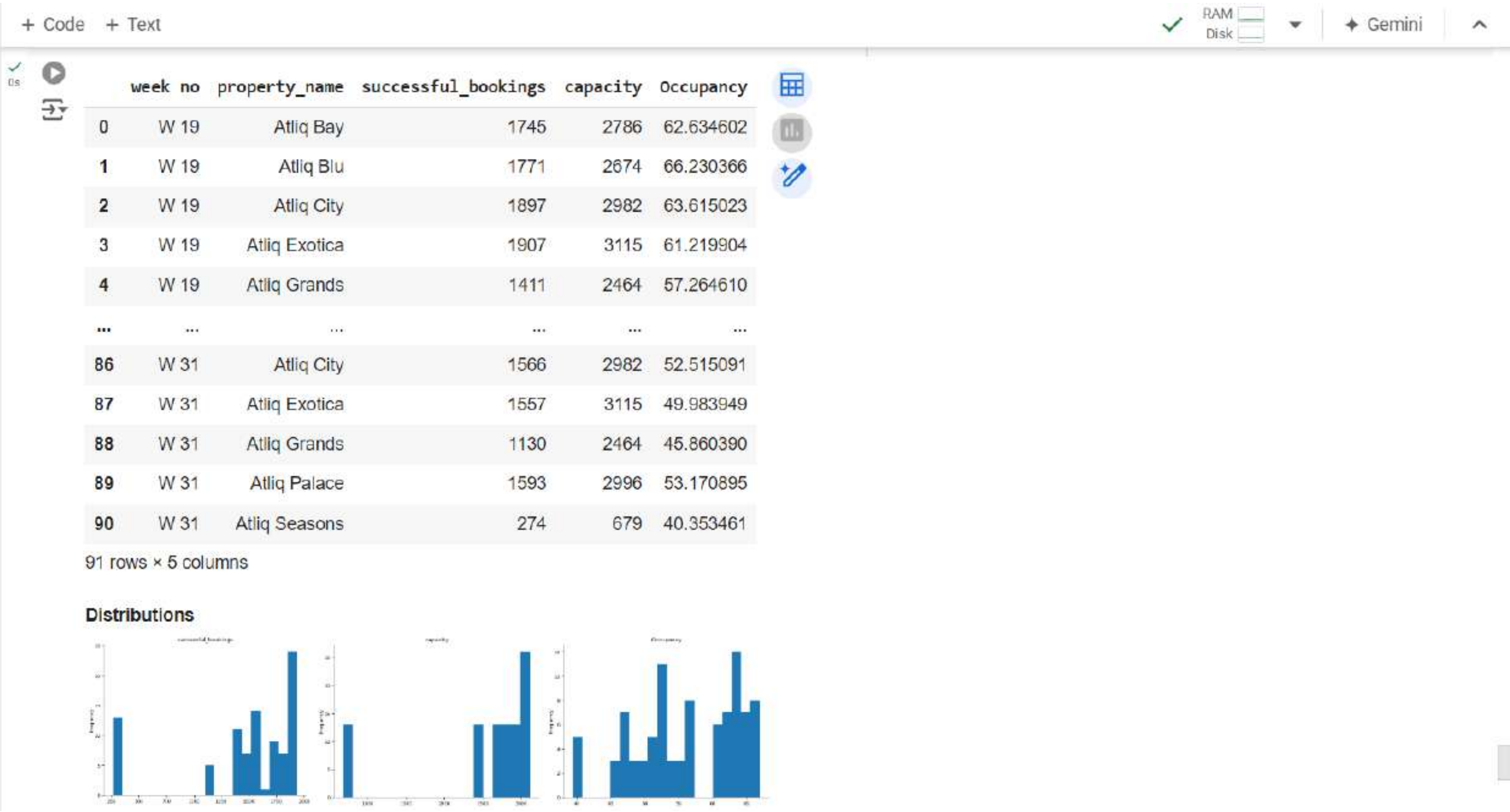
Values



```
z_pivot=z.pivot(index=['Day No','Day Name'],columns='property_name',values='Occupancy')
z_pivot.plot(kind='barh',legend='false')
plt.ylabel('Occupancy Percentage')
plt.title('Daywise Occupancy')
```

```
Text(0.5, 1.0, 'Daywise Occupancy')
```

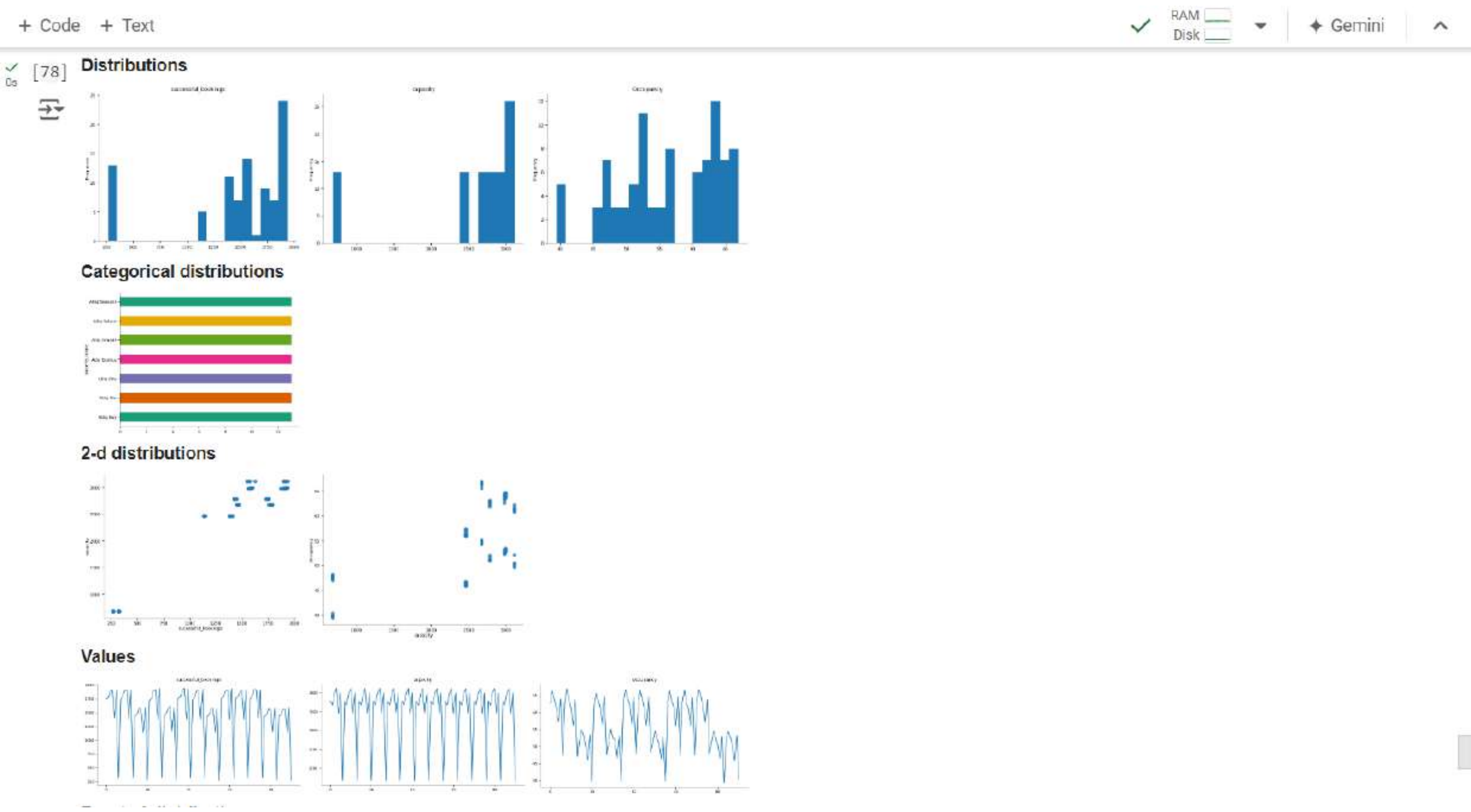
```
[78] weekly_occ=occ_df.groupby(['week no','property_name']).agg({'successful_bookings':'sum','capacity':'sum'})
     weekly_occx=weekly_occ.reset_index()
     weekly_occx['Occupancy']=weekly_occx['successful_bookings']*100/weekly_occx['capacity']
     weekly_occx
```



	week no	property_name	successful_bookings	capacity	Occupancy
0	W 19	Atliq Bay	1745	2786	62.634602
1	W 19	Atliq Blu	1771	2674	66.230366
2	W 19	Atliq City	1897	2982	63.615023
3	W 19	Atliq Exotica	1907	3115	61.219904
4	W 19	Atliq Grands	1411	2464	57.264610
...	...	...	...	...	...
86	W 31	Atliq City	1566	2982	52.515091
87	W 31	Atliq Exotica	1557	3115	49.983949
88	W 31	Atliq Grands	1130	2464	45.860390
89	W 31	Atliq Palace	1593	2996	53.170895
90	W 31	Atliq Seasons	274	679	40.353461

91 rows × 5 columns

Distributions

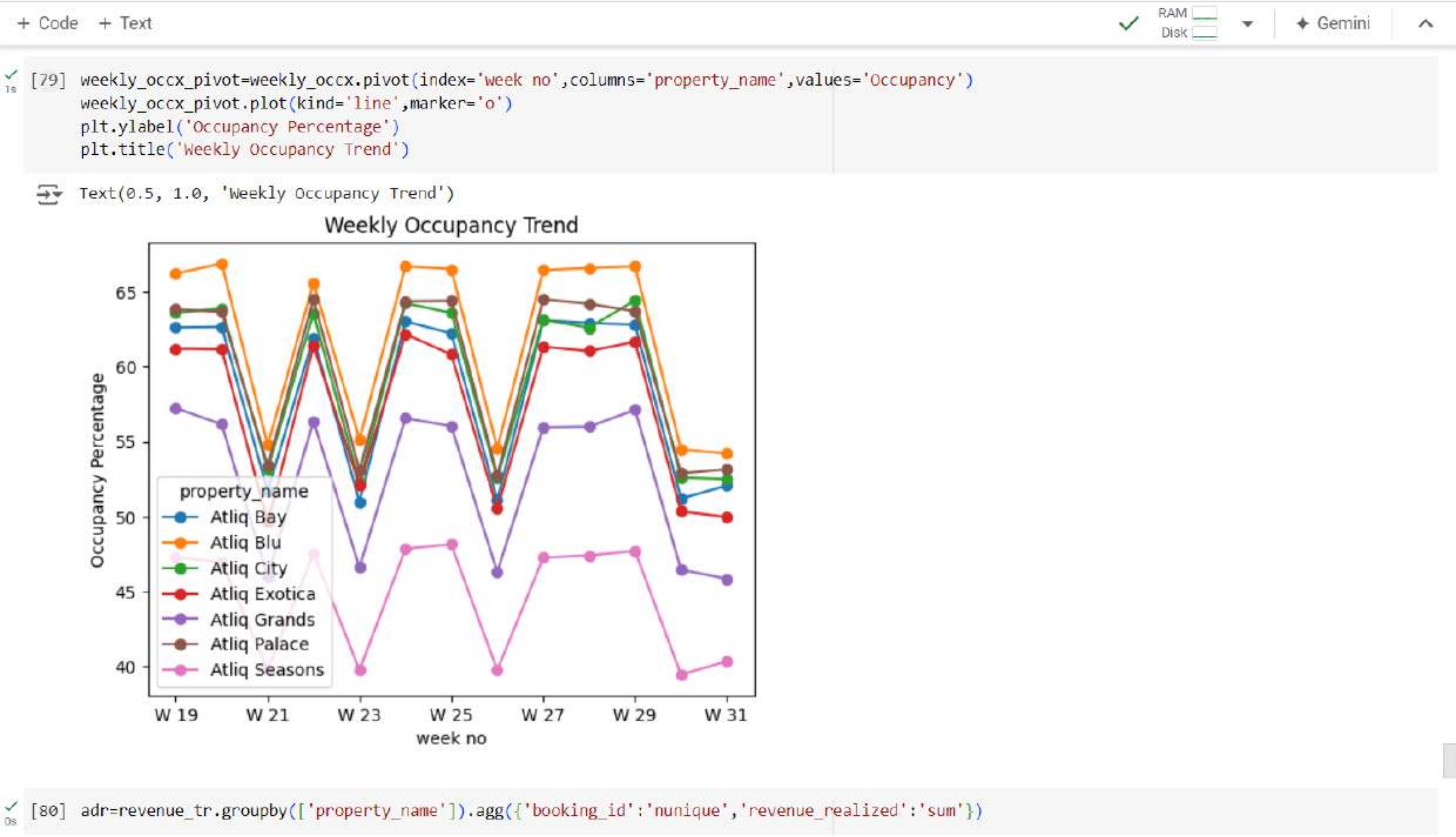


Distributions

Categorical distributions

2-d distributions

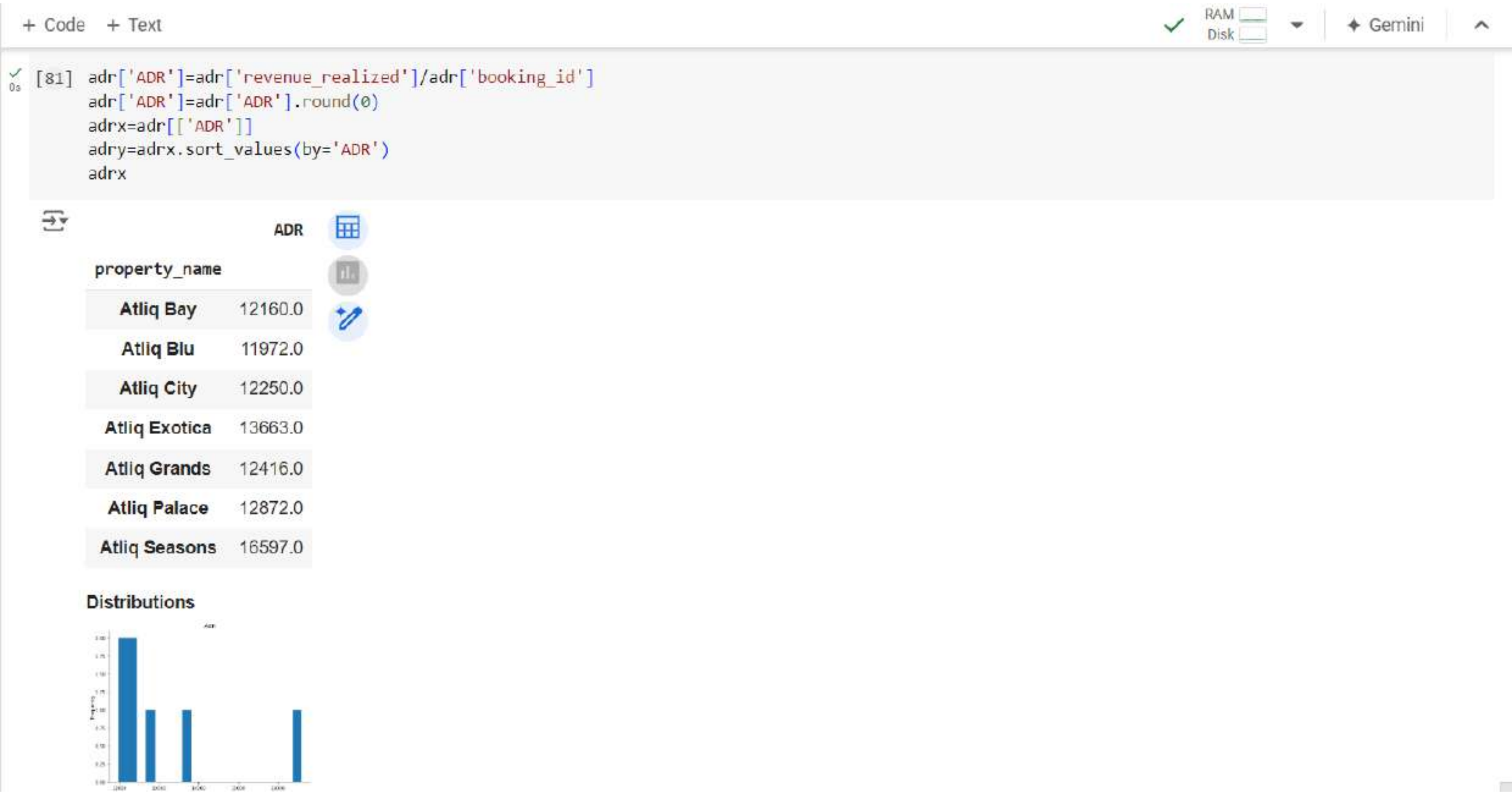
Values



```
[79] weekly_occx_pivot=weekly_occx.pivot(index='week no',columns='property_name',values='Occupancy')
     weekly_occx_pivot.plot(kind='line',marker='o')
     plt.ylabel('Occupancy Percentage')
     plt.title('Weekly Occupancy Trend')
```

```
Text(0.5, 1.0, 'Weekly Occupancy Trend')
```

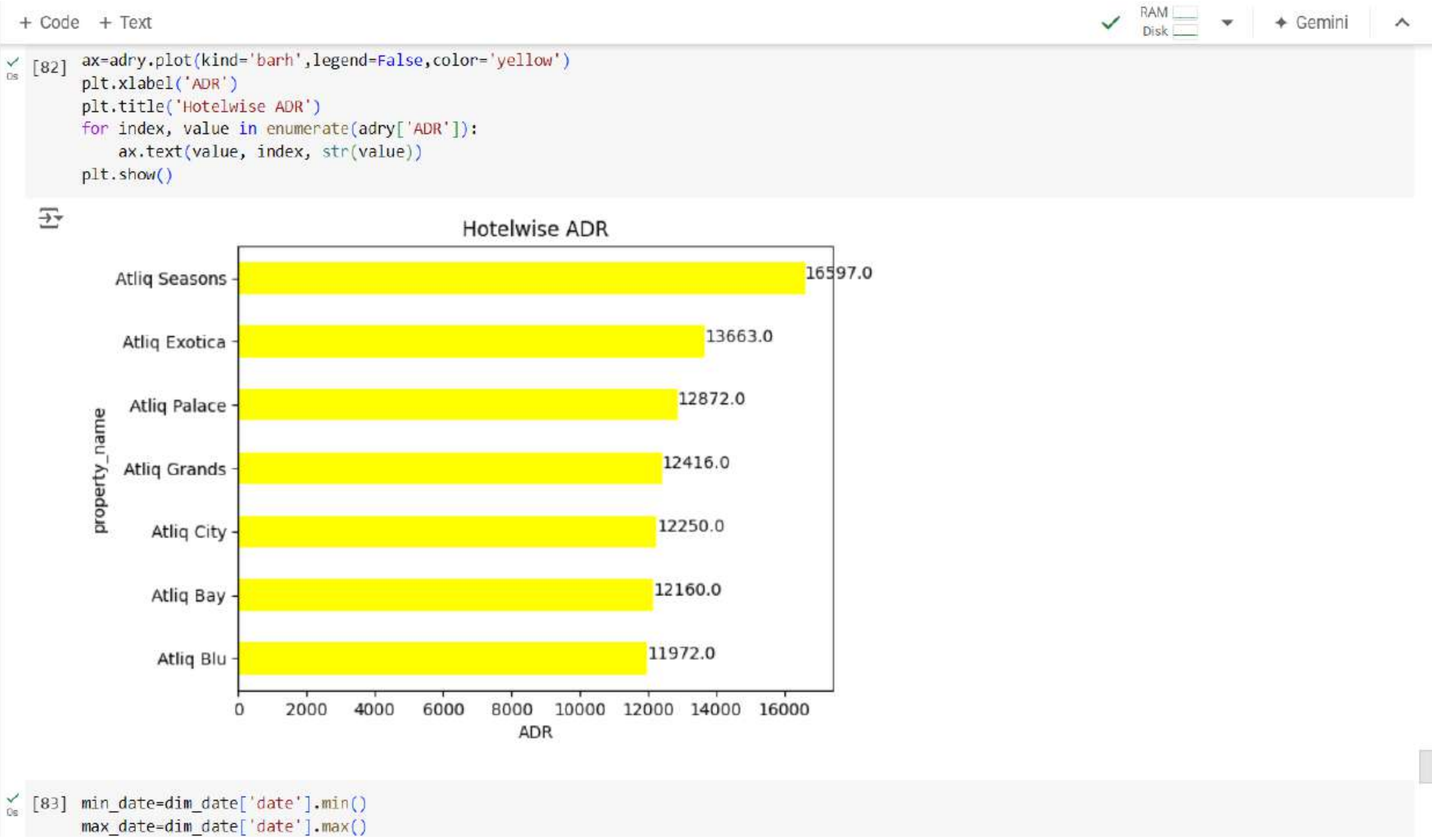
```
[80] adr=revenue_tr.groupby(['property_name']).agg({'booking_id':'nunique','revenue_realized':'sum'})
```



```
[81] adr['ADR']=adr['revenue_realized']/adr['booking_id']
     adr['ADR']=adr['ADR'].round(0)
     adrx=adr[['ADR']]
     adry=adrx.sort_values(by='ADR')
     adrx
```

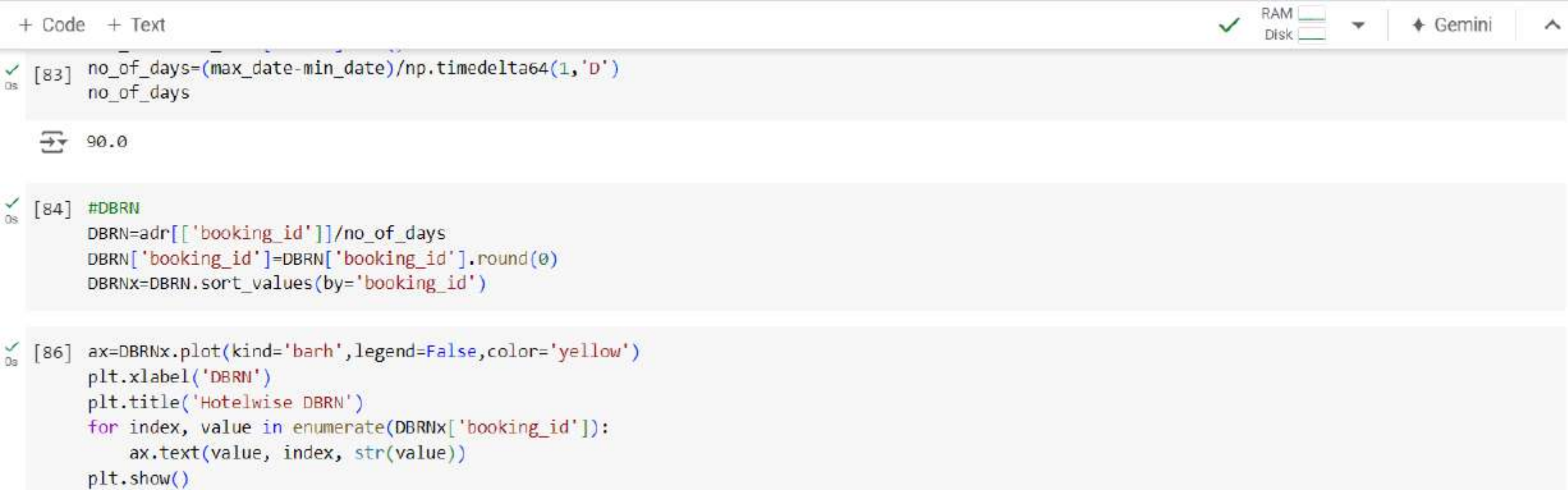
	ADR
property_name	
Atliq Bay	12160.0
Atliq Blu	11972.0
Atliq City	12250.0
Atliq Exotica	13663.0
Atliq Grands	12416.0
Atliq Palace	12872.0
Atliq Seasons	16597.0

Distributions



```
[82] ax=adry.plot(kind='barh',legend=False,color='yellow')
     plt.xlabel('ADR')
     plt.title('Hotelwise ADR')
     for index, value in enumerate(adry['ADR']):
         ax.text(value, index, str(value))
     plt.show()
```

```
[83] min_date=dim_date['date'].min()
     max_date=dim_date['date'].max()
```



[83]

```
no_of_days=(max_date-min_date)/np.timedelta64(1,'D')
no_of_days
```

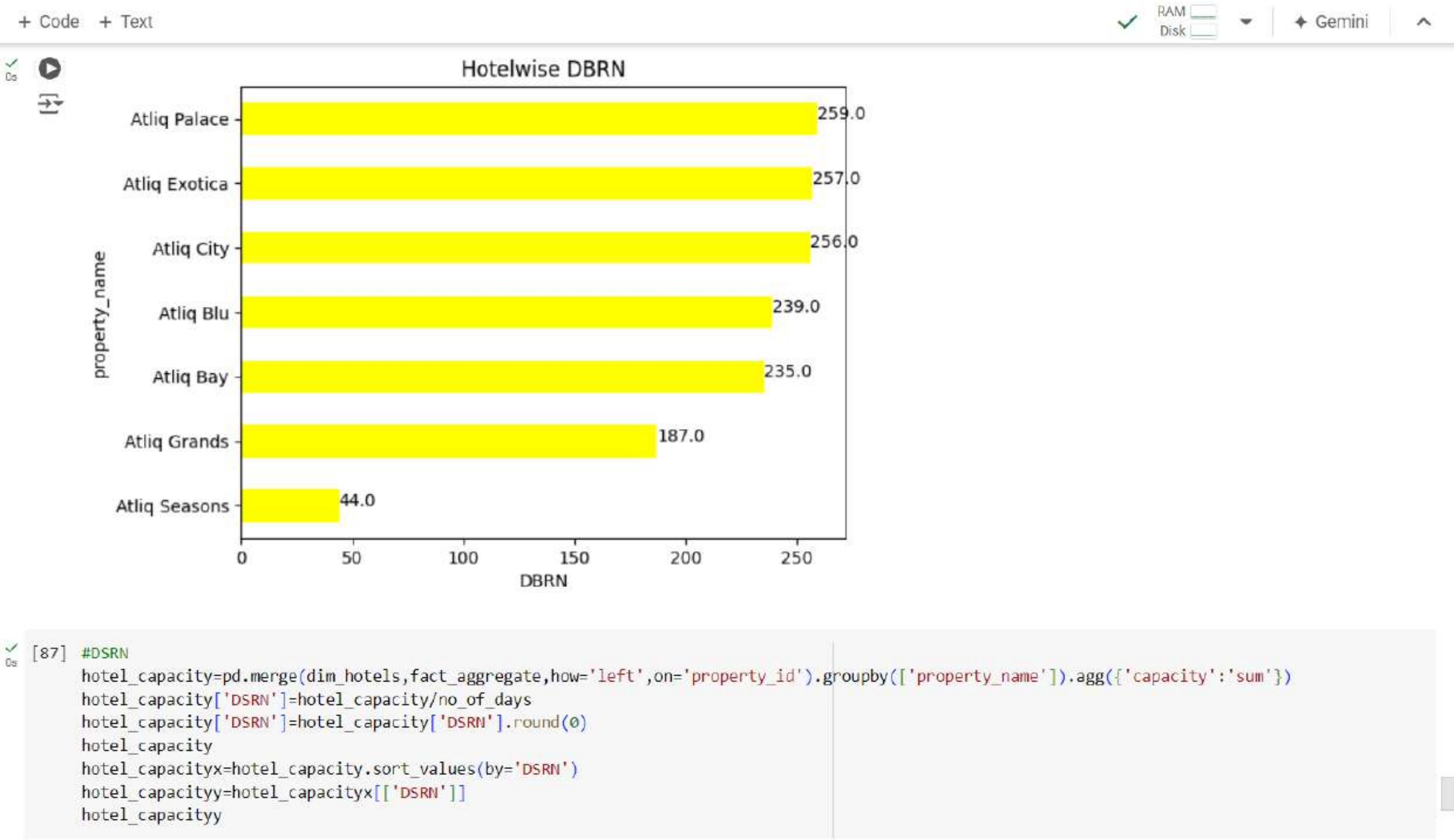
```
90.0
```

[84]

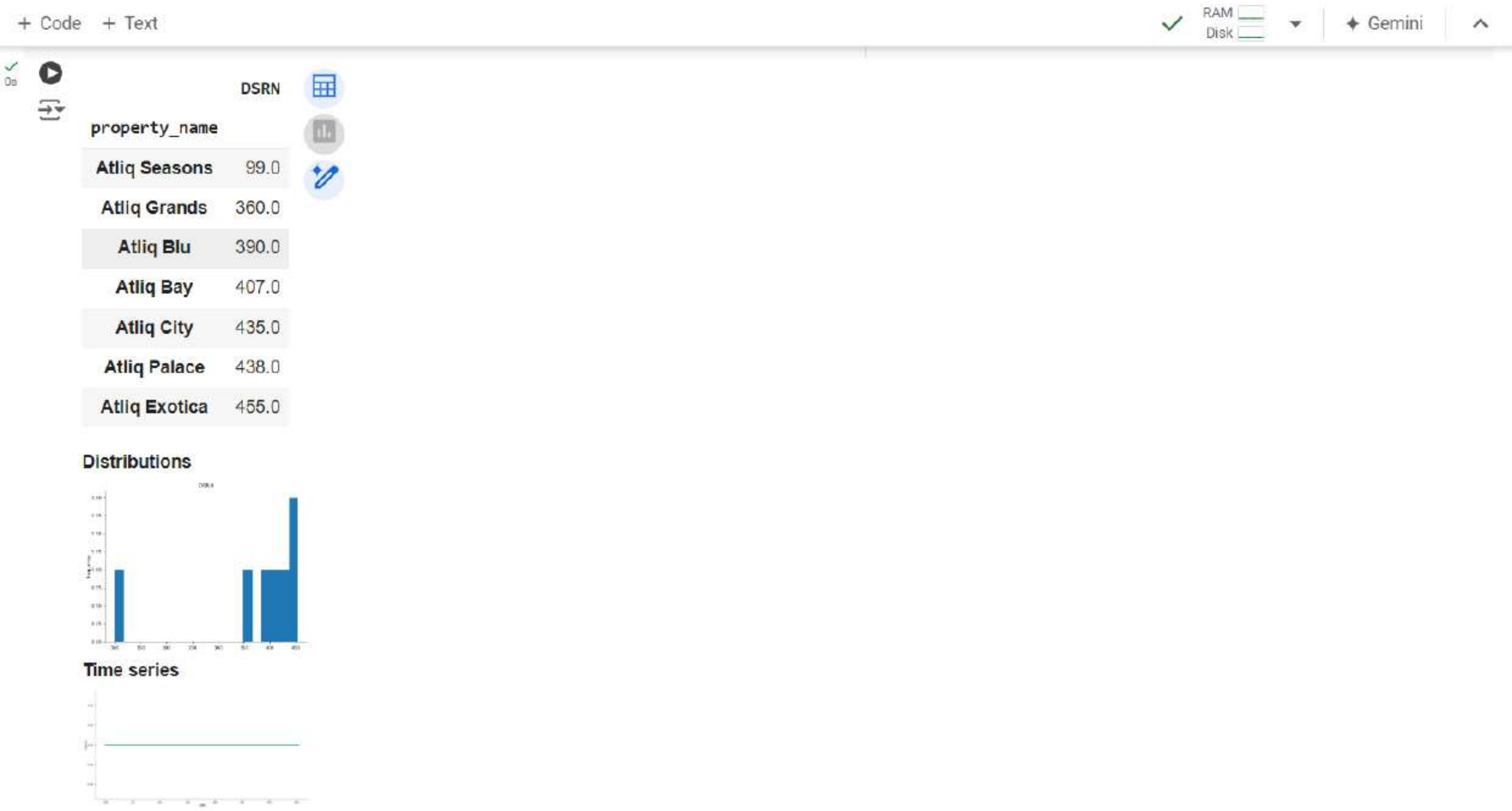
```
#DBRN
DBRN=adr[['booking_id']]/no_of_days
DBRN['booking_id']=DBRN['booking_id'].round(0)
DBRNx=DBRN.sort_values(by='booking_id')
```

[86]

```
ax=DBRNx.plot(kind='barh',legend=False,color='yellow')
plt.xlabel('DBRN')
plt.title('Hotelwise DBRN')
for index, value in enumerate(DBRNx['booking_id']):
    ax.text(value, index, str(value))
plt.show()
```



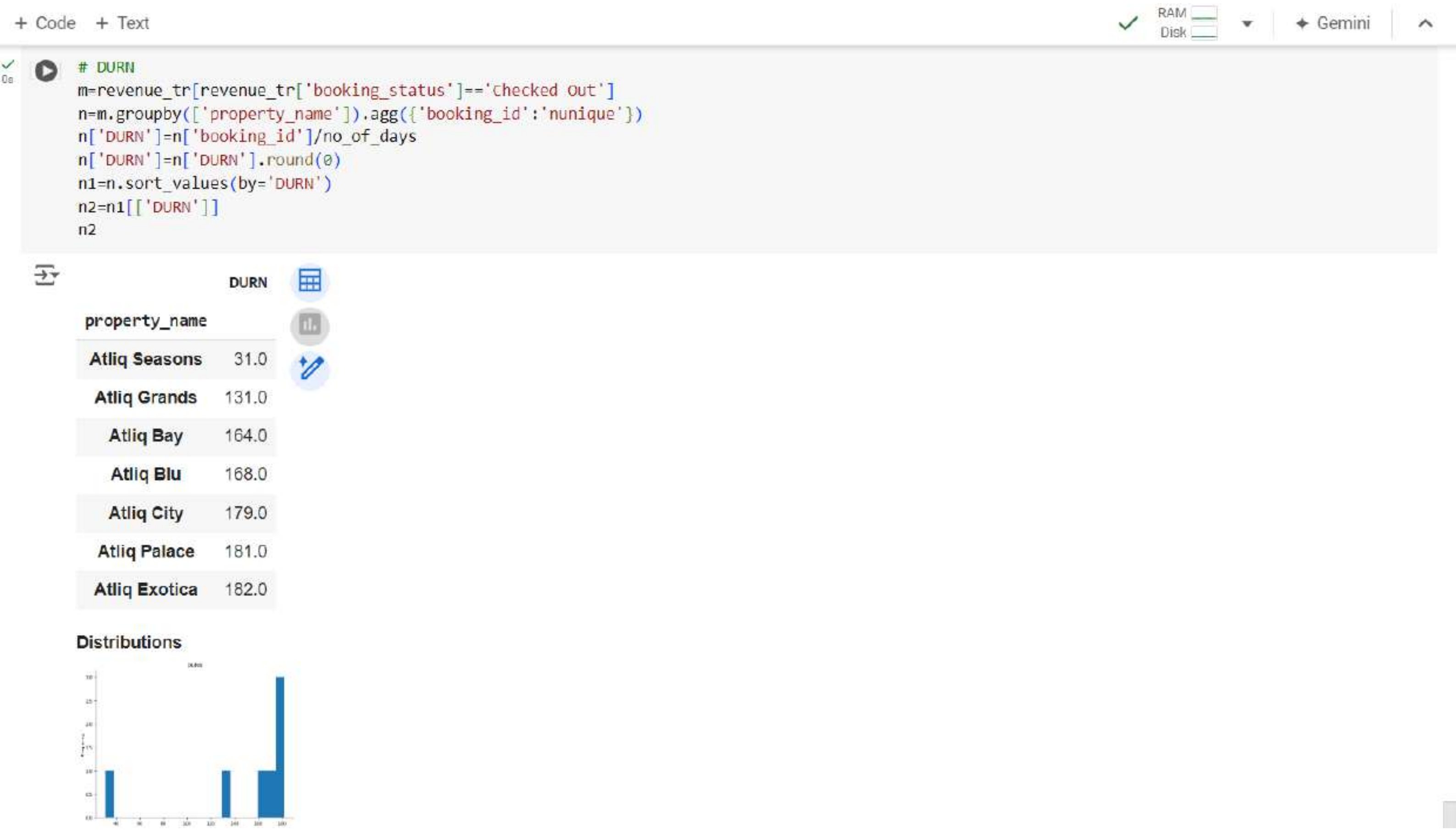
```
#DSRN
hotel_capacity=pd.merge(dim_hotels,fact_aggregate,how='left',on='property_id').groupby(['property_name']).agg({'capacity':'sum'})
hotel_capacity['DSRN']=hotel_capacity/no_of_days
hotel_capacity['DSRN']=hotel_capacity['DSRN'].round(0)
hotel_capacity
hotel_capacityx=hotel_capacity.sort_values(by='DSRN')
hotel_capacityy=hotel_capacityx[['DSRN']]
hotel_capacityy
```



property_name	DSRN
Atliq Seasons	99.0
Atliq Grands	360.0
Atliq Blu	390.0
Atliq Bay	407.0
Atliq City	435.0
Atliq Palace	438.0
Atliq Exotica	455.0

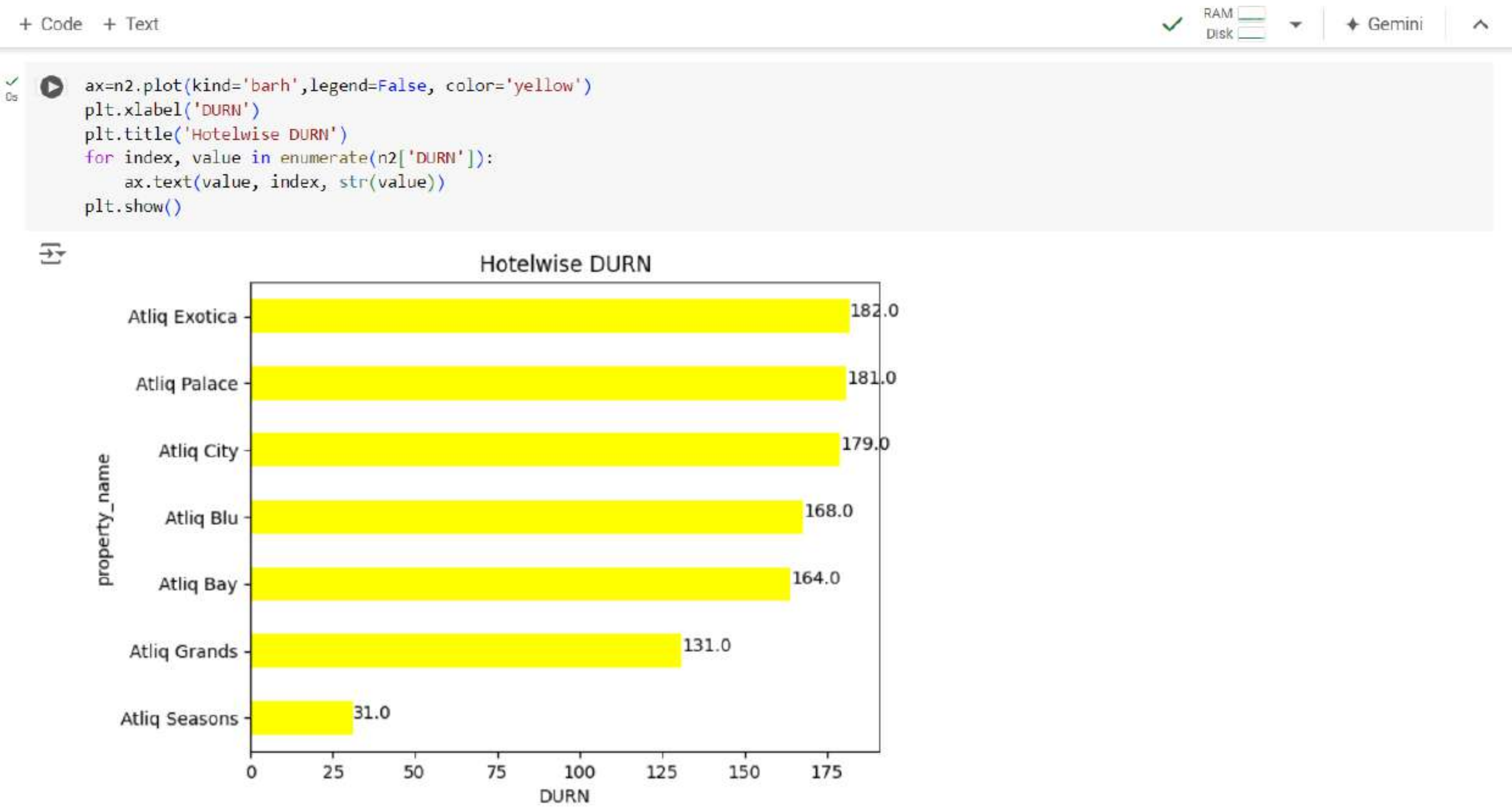
Distributions

Time series



```
# DURN
m=revenue_tr[revenue_tr['booking_status']=='Checked Out']
n=m.groupby(['property_name']).agg({'booking_id':'nunique'})
n['DURN']=n['booking_id']/no_of_days
n['DURN']=n['DURN'].round(0)
n1=n.sort_values(by='DURN')
n2=n1[['DURN']]
n2
```

```
                 DURN
property_name
Atliq Seasons    31.0
Atliq Grands    131.0
Atliq Bay       164.0
Atliq Blu       168.0
Atliq City      179.0
Atliq Palace    181.0
Atliq Exotica   182.0
```



```
ax=n2.plot(kind='barh',legend=False, color='yellow')
plt.xlabel('DURN')
plt.title('Hotelwise DURN')
for index, value in enumerate(n2['DURN']):
    ax.text(value, index, str(value))
plt.show()
```