

Udacity DS project

January 23, 2022

1 Project of Seattle Airbnb

1.1 Questions:

1.1.1 1. What is the average price of the Airbnbs in Seattle?

1.1.2 2. Which variable have the strongest correlation to the price of an Airbnb?

1.1.3 3. Which price range has the most number of Airbnb?

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

1.1.4 Prepare the dataset

```
[2]: calender = pd.read_csv("calendar.csv")
listings = pd.read_csv("listings.csv")
```

```
[3]: calender.head()
```

```
[3]:   listing_id      date available  price
0      241032  2016-01-04          t  $85.00
1      241032  2016-01-05          t  $85.00
2      241032  2016-01-06          f    NaN
3      241032  2016-01-07          f    NaN
4      241032  2016-01-08          f    NaN
```

```
[4]: listings.head()
```

```
[4]:   id      listing_url      scrape_id last_scraped \
0  241032  https://www.airbnb.com/rooms/241032  20160104002432  2016-01-04
1  953595  https://www.airbnb.com/rooms/953595  20160104002432  2016-01-04
2  3308979  https://www.airbnb.com/rooms/3308979  20160104002432  2016-01-04
3  7421966  https://www.airbnb.com/rooms/7421966  20160104002432  2016-01-04
4   278830  https://www.airbnb.com/rooms/278830  20160104002432  2016-01-04
```

```
                                name \
0      Stylish Queen Anne Apartment
```

1 Bright & Airy Queen Anne Apartment
 2 New Modern House-Amazing water view
 3 Queen Anne Chateau
 4 Charming craftsman 3 bdm house

summary \
 0 NaN
 1 Chemically sensitive? We've removed the irrita...
 2 New modern house built in 2013. Spectacular s...
 3 A charming apartment that sits atop Queen Anne...
 4 Cozy family craftman house in beautiful neighb...

space \
 0 Make your self at home in this charming one-be...
 1 Beautiful, hypoallergenic apartment in an extr...
 2 Our house is modern, light and fresh with a wa...
 3 NaN
 4 Cozy family craftman house in beautiful neighb...

description experiences_offered \
 0 Make your self at home in this charming one-be... none
 1 Chemically sensitive? We've removed the irrita... none
 2 New modern house built in 2013. Spectacular s... none
 3 A charming apartment that sits atop Queen Anne... none
 4 Cozy family craftman house in beautiful neighb... none

neighborhood_overview ... review_scores_value \
 0 NaN ... 10.0
 1 Queen Anne is a wonderful, truly functional vi... ... 10.0
 2 Upper Queen Anne is a charming neighborhood fu... ... 10.0
 3 NaN ... NaN
 4 We are in the beautiful neighborhood of Queen 9.0

requires_license license jurisdiction_names instant_bookable \
 0 f NaN WASHINGTON f
 1 f NaN WASHINGTON f
 2 f NaN WASHINGTON f
 3 f NaN WASHINGTON f
 4 f NaN WASHINGTON f

cancellation_policy require_guest_profile_picture \
 0 moderate f
 1 strict t
 2 strict f
 3 flexible f
 4 strict f

	require_guest_phone_verification	calculated_host_listings_count	\
0	f	2	
1	t	6	
2	f	2	
3	f	1	
4	f	1	

	reviews_per_month
0	4.07
1	1.48
2	1.15
3	NaN
4	0.89

[5 rows x 92 columns]

```
[5]: df1 = pd.merge(calender, listings, left_on='listing_id', right_on='id',
    ↪how='left').drop('id', axis=1)
```

```
[6]: df1 = df1.fillna(-1)
```

```
[7]: df1.head()
```

```
[7]: listing_id      date available price_x \
0      241032  2016-01-04      t  $85.00
1      241032  2016-01-05      t  $85.00
2      241032  2016-01-06      f    -1
3      241032  2016-01-07      f    -1
4      241032  2016-01-08      f    -1
```

	listing_url	scrape_id	last_scraped	\
0	https://www.airbnb.com/rooms/241032	20160104002432	2016-01-04	
1	https://www.airbnb.com/rooms/241032	20160104002432	2016-01-04	
2	https://www.airbnb.com/rooms/241032	20160104002432	2016-01-04	
3	https://www.airbnb.com/rooms/241032	20160104002432	2016-01-04	
4	https://www.airbnb.com/rooms/241032	20160104002432	2016-01-04	

	name	summary	\
0	Stylish Queen Anne Apartment	-1	
1	Stylish Queen Anne Apartment	-1	
2	Stylish Queen Anne Apartment	-1	
3	Stylish Queen Anne Apartment	-1	
4	Stylish Queen Anne Apartment	-1	

	space	... review_scores_value	\
0	Make your self at home in this charming one-be...	...	10.0
1	Make your self at home in this charming one-be...	...	10.0

2	Make your self at home in this charming one-be...	...	10.0
3	Make your self at home in this charming one-be...	...	10.0
4	Make your self at home in this charming one-be...	...	10.0

	requires_license	license	jurisdiction_names	instant_bookable	\
0	f	-1.0	WASHINGTON	f	
1	f	-1.0	WASHINGTON	f	
2	f	-1.0	WASHINGTON	f	
3	f	-1.0	WASHINGTON	f	
4	f	-1.0	WASHINGTON	f	

	cancellation_policy	require_guest_profile_picture	\
0	moderate	f	
1	moderate	f	
2	moderate	f	
3	moderate	f	
4	moderate	f	

	require_guest_phone_verification	calculated_host_listings_count	\
0	f	2	
1	f	2	
2	f	2	
3	f	2	
4	f	2	

	reviews_per_month
0	4.07
1	4.07
2	4.07
3	4.07
4	4.07

[5 rows x 95 columns]

1.2 1. What is the average price of the Airbnbs in Seattle?

```
[8]: import plotly.offline as py
import plotly.express as px
import plotly.graph_objects as go
import plotly.offline as pyo
from plotly.subplots import make_subplots
import seaborn as sns
from statistics import mean
```

/Users/mapengfei/opt/anaconda3/lib/python3.7/site-packages/statsmodels/tools/_testing.py:19: FutureWarning:

pandas.util.testing is deprecated. Use the functions in the public API at pandas.testing instead.

```
[9]: prices = [i for i in df1['price_x'] if not -1]

for i in list(df1['price_y']):
    if i != -1:
        prices.append(i)
```

```
[10]: price = []

for i in prices:
    i = i[:-2]
    i = ''.join(e for e in i if e.isalnum())
    price.append(int(i))
```

```
[11]: average = mean(price)
```

```
[12]: average
```

```
[12]: 127.97616553169199
```

```
[13]: trace = go.Box(y=price, name='Price')

data = [trace]

layout = go.Layout(title='Boxplot of Price', hovermode='x')

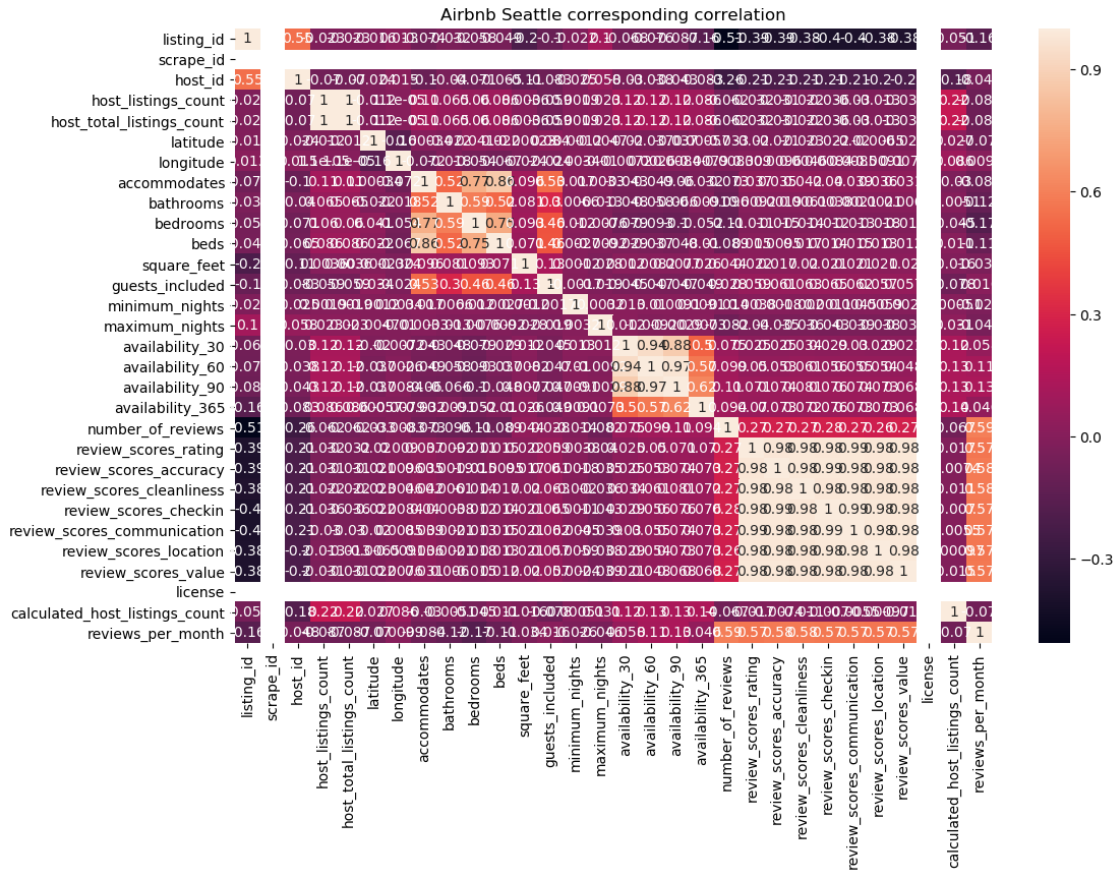
fig = go.Figure(data=data, layout=layout)

fig.show()
```

It is easy to see that the average price of Airbnb in Seattle is \$127.98

1.3 2. Which variable have the strongest correlation to the price of an Airbnb?

```
[14]: fig=plt.figure(figsize=(12,8), dpi= 100, facecolor='w', edgecolor='k')
MO = sns.heatmap(df1.corr(), annot = True).set_title('Airbnb Seattle_
→corresponding correlation')
```



Based on the color of the heatmap. It is easy to see that all variables related to reviews and availability. `host_listings_count` have very strong correlation to the price of an Airbnb.

1.4 3. Which price range has the most number of Airbnb?

```
[15]: fig = px.histogram(price)
fig.show()
```

So it is clearly to see that most of the price of airbnbs are between 35 to 175 and 150 has the most number of airbnbs

```
[ ]:
```