```
# We've loaded your first package for you! You can add as many cells as you
need.
import numpy as np
import pandas as pd
import datetime as dt

#1 Importing the data
prices = pd.read_csv("data/airbnb_price.csv")
xls=pd.ExcelFile("data/airbnb_room_type.xlsx")
room_types = xls.parse(0)
reviews = pd.read_csv("data/airbnb_last_review.tsv", sep='\t')

#2 Cleaning the price column
prices["price"] = prices["price"].str.replace(" dollars", "")
prices["price"] = pd.to_numeric(prices["price"])

#3 Calculating average price
zero_listings = prices["price"] == 0
prices = prices.drop(prices[zero_listings].index)
avg_price = round(prices["price"].mean(), 2)

#4 Comparing costs to the private rental market
prices["prices_per_month"]=prices["price"]*365 / 12
average_price_per_month=round(prices["prices_per_month"].mean(), 2)
difference= round(average_price_per_month-3100,2)

#5 Cleaning the room_type column
room_types["room_type"]=room_types["room_type"].str.lower()
room_types["room_type"]=room_types["room_type"].astype("category")
room_frequencies = room_types["room_type"].value_counts()

#6 What timeframe are we working with?
reviews["last_review"] = pd.to_datetime(reviews["last_review"])
first_reviewed=reviews["last_review"].dt.date.min()
last_reviewed=reviews["last_review"].dt.date.max()

#7 Joining the DataFrames
rooms_and_prices = prices.merge(room_types, how="outer", on="listing_id")
airbnb_merged=rooms_and_prices.merge(reviews, how="outer", on="listing_id")
airbnb_merged = airbnb_merged.dropna()

#8 Analyzing listing prices by NYC borough
airbnb_merged["borough"] = airbnb_merged["nbhood_full"].str.split(",").str[0]
boroughs =
airbnb_merged.groupby("borough")["price"].agg(["sum","mean","median","count"])
boroughs=boroughs.round(2).sort_values("mean",ascending=False)

#9 Price range by bourough
```

```
label_names=["Budget","Average","Expensive","Extravagant"]
ranges=[0,69,175,350,np.inf]
airbnb_merged["price_range"] = pd.cut(airbnb_merged["price"], bins=ranges,
labels=label_names)
prices_by_borough = airbnb_merged.groupby(["borough",
"price_range"])["price_range"].count()

#10 Storing the final result
airbnb_analysis={'avg_price':avg_price,
                 'average_price_per_month':average_price_per_month,
                 'difference':difference,
                 'room_frequencies':room_frequencies,
                'first_reviewed':first_reviewed,
                 'last_reviewed':last_reviewed,
                 'prices_by_borough':prices_by_borough}
print(airbnb_analysis)
```

OUTPUT:

```
shared room          587
Name: room_type, dtype: int64, 'first_reviewed': datetime.date(2019, 1, 1),
'last_reviewed': datetime.date(2019, 7, 9), 'prices_by_borough':
borough        price_range
Bronx          Budget          381
               Average         285
               Expensive        25
               Extravagant       5
Brooklyn       Budget         3194
               Average        5532
               Expensive      1466
               Extravagant     259
Manhattan      Budget         1148
               Average        5285
               Expensive      3072
               Extravagant     810
Queens         Budget         1631
               Average        1505
               Expensive       291
               Extravagant      28
Staten Island  Budget          124
               Average         123
               Expensive        20
               Extravagant       0
Name: price_range, dtype: int64}
```