Low Level Design (LLD)

Travel Data Analysis

(AirBnB Data Analysis)

Written By / Author	Lokesh Attarde
Document Version	LLD-V1.0
Last Revised Date	15/10/2021

Document Version Control

Date Issued	Version	Description	Author
27/05/2022	LLD-V1.0	First Version of Complete LLD	Roshan

Contents

Abstract

Airbnb is an American company that facilitates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. It basically connecting travelers with local hosts who want to rent out their homes with people who are looking for accommodations in that locality. On the other hand, this platform enables host to list their available space and earn extra income in the form of rent and it also enables travelers to book unique homestays from local hosts, saving them money and giving them a chance to interact with locals.

In the world of rising new technology and innovation, Travel industry is advancing with the role of Data Science and Analytics. Data analysis can help them to understand their business in a quiet different manner and helps to improve the quality of the service by identifying the weak areas of the business. This study demonstrates the how different analysis help out to make better business decisions and help analyze customer trends and satisfaction, which can lead to new and better products and services. Different analysis performed such as Exploratory Data Analysis and Descriptive Analysis on variety of use cases to get the key insights from this data based on which business decisions will be taken.

1 Introduction

1.1 Why this Low-Level design document?

The purpose of this LLD or a Low-Level Design (LLD) document is to give the internal logical design of the actual program code for Airbnb Data Analysis project. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document. This document is intended for both the stakeholders and the developers of this project and will be proposed to the higher management for its approval.

The main objective of the project is to analyse the various aspects with different use cases which covers many aspects of airbnb listings. It helps in not only understanding the meaningful relationships between attributes but it also allows us to do our own research and come-up with our findings.

1.2 Scope

Low-level design (LLD) is a component-level design process that follows a step-by step refinement process. This process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

This study demonstrates the how different analysis help out to make better business decisions and help analyse customer trends and satisfaction, which can lead to new and better products and services.

1.3 Constraints

The analysis must be user friendly, code must be neat & clean, EDA must be automated as much as possible because it will save huge amount of time. Moreover, users should not be required to have any of the coding knowledge as the insights they are looking for are mentioned in-detail with respective visuals.

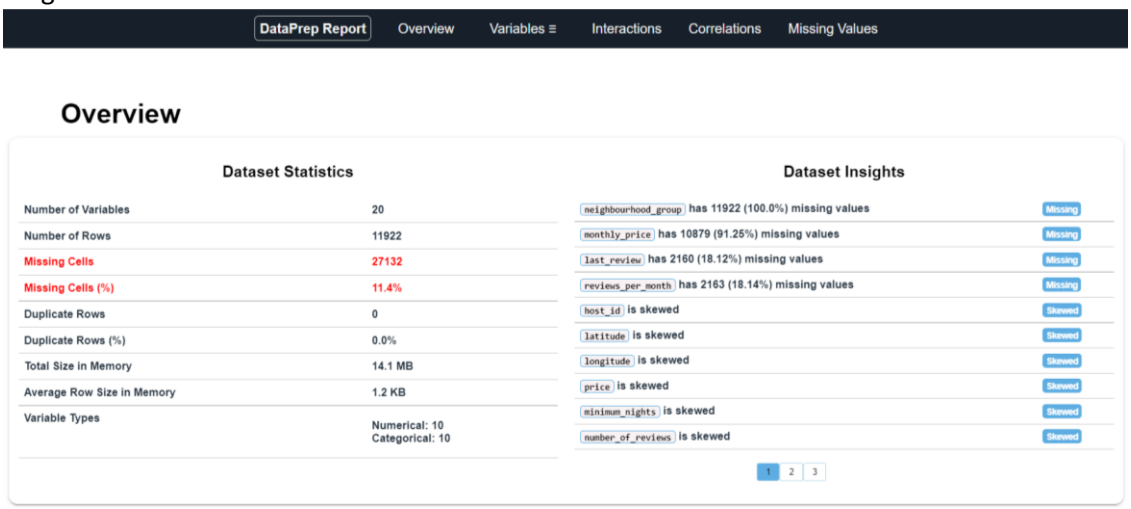
2 Technical Specifications

2.1 Listings Dataset -

id	name	host_id	host_nam	neighbourl	neighbourl	city	latitude	longitude	property_type	room_type	amenities	price	monthly_price	minimum_	number_	last_review	reviews_p	calculated	availability_365
6	Large Craf	29	Sara		North Hills	San Diego	32.75399	-117.13	House	Entire home/apt	{TV,Internet,Wifi,"Air	295		2	149	10/7/2018	1.18	1	20
5570	Ocean fro	8435	Jef Karchin'S MISSION		Mission Be	San Diego	32.7843	-117.253	Condominium	Entire home/apt	{TV,"Cable TV",Intern	2050		3	0			2	328
8095	Sunset Clif	270	Marin		Ocean Bea	San Diego	32.73517	-117.244	Guesthouse	Entire home/apt	{Internet,Wifi,Kitcher	84		2	178	11/2/2015	1.56	1	0
39516	Art Studio	169649	Chris And Jean		North Hills	San Diego	32.73188	-117.119	Tiny house	Entire home/apt	{Internet,Wifi,Pool,Ki	69		3	338	11/7/2018	3.33	1	273
45429	OB cottag	197919	Melissa		Loma Port	San Diego	32.74877	-117.229	House	Entire home/apt	{TV,Internet,Wifi,Kitc	160		2	6	7/28/2014	0.06	1	363
54001	La Jolla Pe	252692	Marsha		La Jolla	La Jolla	32.8119	-117.27	House	Private room	{TV,"Cable TV",Wifi,K	105	$2,150.00	2	174	10/28/2018	1.82	4	304
62274	charming,	302986	Isabel		Pacific Bea	San Diego	32.80734	-117.242	Guesthouse	Entire home/apt	{Wifi,"Air conditionin	68	$2,500.00	2	460	10/30/2018	4.75	2	249
62949	Pacific Bea	21532	Lisa		Pacific Bea	San Diego	32.80961	-117.231	Guest suite	Entire home/apt	{TV,"Cable TV",Wifi,"	75		1	498	11/10/2018	6.56	2	119
67441	Canyon ap	332552	Peter		College Ar	San Diego	32.76365	-117.077	Apartment	Entire home/apt	{TV,"Cable TV",Intern	90	$1,450.00	3	3	7/13/2015	0.07	5	285
69385	Mission Be	347408	Kimberly		Mission Be	San Diego	32.78646	-117.254	House	Entire home/apt	{TV,"Cable TV",Intern	250	$5,500.00	3	162	10/29/2018	1.72	1	296
75668	San Diego	401936	Johnathon		North Hills	San Diego	32.7473	-117.144	Condominium	Private room	{TV,"Cable TV",Intern	50	$1,050.00	2	168	10/28/2018	1.89	2	204
79300	San Diego	401936	Johnathon		North Hills	San Diego	32.74594	-117.144	Condominium	Private room	{TV,"Cable TV",Intern	59	$1,100.00	2	148	10/28/2018	1.59	2	230
103417	La Jolla Re	252692	Marsha		La Jolla	La Jolla	32.8128	-117.267	House	Private room	{TV,"Cable TV",Intern	125	$1,895.00	2	145	9/30/2018	1.64	4	79
103420	La Jolla Se	252692	Marsha		La Jolla	La Jolla	32.81253	-117.268	House	Private room	{TV,"Cable TV",Wifi,K	115	$2,400.00	2	113	11/12/2018	1.32	4	160
103720	Gaslamp R	507423	Ann		Midtown	San Diego	32.74886	-117.165	Bed and breakf	Private room	{TV,"Cable TV",Intern	119		1	44	8/5/2018	0.49	6	148
103841	Peaceful &	541128	Flo		North City	San Diego	32.95923	-117.233	Townhouse	Private room	{TV,Internet,Wifi,Poc	59	$1,100.00	2	86	11/13/2018	0.95	1	252
108349	3Br Townh	546519	Juliana		Pacific Bea	San Diego	32.7966	-117.253	Townhouse	Entire home/apt	{TV,"Cable TV",Intern	209		2	115	10/23/2018	1.31	2	325
117896	La Jolla Be	470998	Miriam		La Jolla	San Diego	32.8107	-117.261	House	Private room	{Internet,Wifi,"Air co	89		1	163	6/27/2016	1.86	3	0
119063	Cheery Yel	601369	Yellow Beach House		Pacific Bea	San Diego	32.79757	-117.247	House	Entire home/apt	{TV,"Cable TV",Intern	178	$3,000.00	3	16	9/23/2018	0.18	1	209
123990	Laid-Back	479932	Amber		Rancho Pe	San Diego	32.99091	-117.167	House	Entire home/apt	{TV,"Cable TV",Wifi,"	475		5	8	7/12/2018	0.09	1	363
126344	3Br House	546519	Juliana		Pacific Bea	San Diego	32.80572	-117.259	House	Entire home/apt	{TV,"Cable TV",Intern	221		2	126	10/21/2018	1.47	2	253
132966	Water Vie	555739	Rosana		Pacific Bea	San Diego	32.80855	-117.258	Apartment	Entire home/apt	{TV,"Cable TV",Intern	70	$1,800.00	30	18	8/31/2018	0.25	1	153
141523	Private spc	686496	Matt And Toni		Ocean Bea	San Diego	32.75068	-117.247	Guest suite	Entire home/apt	{TV,Internet,Wifi,"Pai	115	$2,900.00	2	295	10/28/2018	3.33	2	36
146886	Downtowr	710122	Kelly		Grant Hill	San Diego	32.70803	-117.143	Apartment	Entire home/apt	{TV,"Cable TV",Wifi,"	79	$1,999.00	7	13	4/18/2018	0.16	11	153
149548	Mission Ba	721326	Olivia		Bay Park	San Diego	32.80157	-117.201	House	Private room	{TV,Internet,Wifi,"Air	49		3	156	10/10/2018	1.76	2	167
161818	Kensingtor	773869	Steve And Kyle(She'S		Midtown	San Diego	32.74183	-117.179	Townhouse	Private room	{TV,"Cable TV",Intern	79	$1,150.00	28	50	10/9/2017	0.57	1	362

2.1.1 Listings Dataset Overview –

The Listings dataset consists of a table with 11922 records and 20 features. Features are distributed as 10 Continuous features and 10 Categorical features. There are a total 11.4% of records having Missing values.

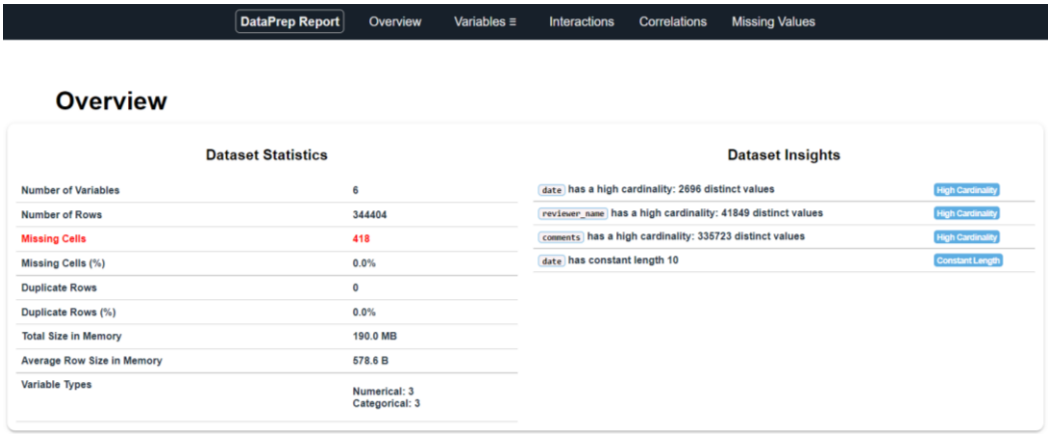


2.2 Reviews Dataset -

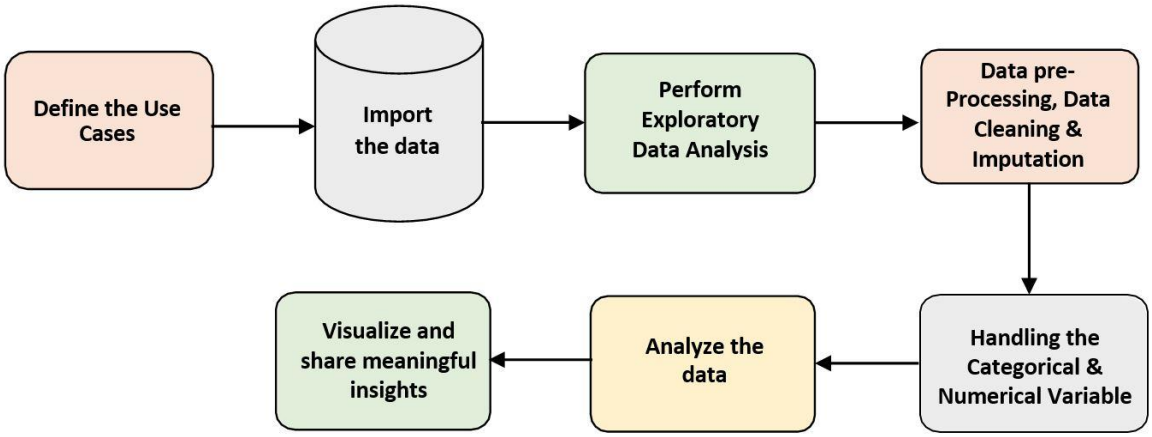
listing_id	id	date	reviewer_id	reviewer_name	comments
6	8	6/22/2008	415	Terrence	Sara is an awesome hostess. She was very considerate, making me feel right at home. Her suggestions of local well-known and up-and-coming hotspots were also very helpful. Definitely a pc
6	4383	6/22/2009	3762	Christine	My stay at â€œCraftsmenâ€ was wonderful! Sara was a lovely hostess, has a cute and stylish home, and her dog is super adorable. She is not kidding, itâ€™s the most comfortable futon ever!
6	1722383	7/16/2012	1864778	Hieu	This was
6	5595071	7/7/2013	117966	Daryna	Sara was
6	5631156	7/8/2013	6563815	Mathew	Sara was such an amazing hostess! My girlfriend and I stayed only one night but felt welcome as soon as we walked through the door. Our room was of high end hotel quality with all the comfe
6	5916942	7/22/2013	7108557	Benjamin	Staying with Sara was very easy and convenient. Access to the unit was simple and the bedroom and bathroom themselves were clean and well maintained. The location was very convenient
6	6090678	7/29/2013	7486194	Jim	Sara and Weston's place is great. Clean, comfortable and centrally located. You must be a dog lover though because Bodie will love on you! :-)
6	6444835	8/12/2013	7938299	Tamara	We enjoyed our stay at Sara's. It was a quick trip, so we weren't there long, but Sara provided us with great information on places to dine and how to get around. Our room was nice and the b
6	6676387	8/20/2013	3169037	Danny	San Diego, with its beaches, weather and great people made my trip a very enjoyable one and Sara and Weston fit into the category of great people easily. They are warm and friendly and are
6	7079682	9/5/2013	3195991	Alejandro	Sara, Weston, and Bodie are great hosts and wonderful folks. My stay was great and more than I ever expected. I would highly recommend this place to anyone else looking for a cool house w
6	7491046	9/22/2013	2372429	Jose & Anabell	My wife and I enjoyed our stay in Sara's house for our 2 nights there. Sara was very welcoming and gave us great advise on what to do. The location is in a really cool neighborhood, with cool l
6	7728349	9/30/2013	4079966	Chris&Tomi	We really enjoyed the neighborhood here. Lots of fun restaurants and bars within close distance but far enough for a peaceful night when you're ready to sleep! Comfy and clean house too! E
6	8036506	10/13/2013	7578342	Zae	Nice person,Nice smile and DOPE ass SPOT!!! will be coming back for sure,and her dog is chill as hell :)
6	8398021	10/28/2013	9603814	Martha	We are local county dwellers, but needed a place closer to North Park to stay because of a Halloween party. We were able to make arrangements with Sara in a very last minute manner, whi
6	8424381	10/29/2013	9553835	Joel	Overall a great experience. It is a really nice, clean room and bathroom and they were great host. In terms of space and ammenities, it felt more like a hotel than someones house. The house
6	8767683	11/17/2013	9275135	Alice	Sara is a
6	10769917	3/6/2014	331740	Shanna	Although Sara was not there, her kind and friendly boyfriend was. My stay was comfortable and I had everything that I needed. But, if big, I mean BIG dogs are not your thing, then be alerted th
6	11118343	3/21/2014	2199062	Ryan	I arrived
6	11135703	3/22/2014	12188990	Natalia	Sara's place is great, just as described! The house is a beautiful craftsman in an up-and-coming North Park neighborhood of San Diego. Shops, restaurants, dance and yoga studios all a short wa
6	11337222	3/30/2014	13324231	Hsin	Had an awesome stay here! Sara was a lifesaver and even let me book last-second for an extra 2 days. There's great food here - all within walking distance - and parking isn't too bad. Her dogs
6	11633084	4/10/2014	10684047	Stephan	Sara really makes you feel at home in her cosy house. She is a great host and treats you like an old friend. The neighbourhood is perfect for a going out at night and a very good burger place is j
6	11925494	4/19/2014	5814525	Dina	Sara was absolutely lovely! Me and my friend Charlotte stayed with her for 3 days and really enjoyed the whole experience. Sara's place is perfectly located- near some cool bars and restaura
6	12262944	4/27/2014	7157690	Mynor	Finally completed my first airbnb stay. Im excited to have stayed at Sara's. Very cool place to crash. Visit SD for a work/play trip and had to book last min. Sara was great, room was nice. True
6	12323806	4/28/2014	14731877	Jiyon	Great location. Comfy bed. Cute patio. Bottle of wine and candles a nice touch. Two Irish Wolfhounds greeted us at the door. Don't be intimidated by their size, they are gentle and well beh
6	12592507	5/5/2014	14861322	Rebecca	This was my first Airbnb experience and I could not ask for a better one! Sara was such a wonderful host! From the second she greeted me and my boyfriend at the door she made us feel welc
6	13045862	5/18/2014	202002	Monica	Sara is very Sweet and I am very thankful she took me in for my short stay. Very cozy room and very nice location , convenient, close to everything.

2.2.1 Reviews Dataset Overview –

Following EDA report illustrate that this Reviews dataset consists of a table with 344404 records and 6 features. Features are distributed as 3 Continuous features and 3 Categorical features. There are an only 418 cells having Missing values.



3 Architecture



3.1 Architecture Description –

3.1.1 Data Description –

As we have seen earlier, in our listings dataset, we have around 1.19 Lacs of records with 20 different features. Features are distributed as 10 Continuous features and 10 Categorical features and in our reviews dataset, we have around 3.44 Lacs of records with 6 different features among them there are 3 Continuous features and 3 Categorical features. These datasets are given in the form of Comma Separated Value (.csv) format.

3.1.2 Define the Use Cases –

At this stage, based on the given dataset and business problems we have defined the several Use Cases to perform the analysis on and this will definitely help out get the key insights from this data based on which business decisions will be taken. Furthermore, It helps in not only understanding the meaningful relationships between attributes but it also allows us to do our own research and come-up with our findings.

3.1.3 Import the Dataset –

As we have received the dataset in the form of Comma Separated Value (.csv) format, therefore we can import the same using Pandas read_csv( ) function.

Reading Data

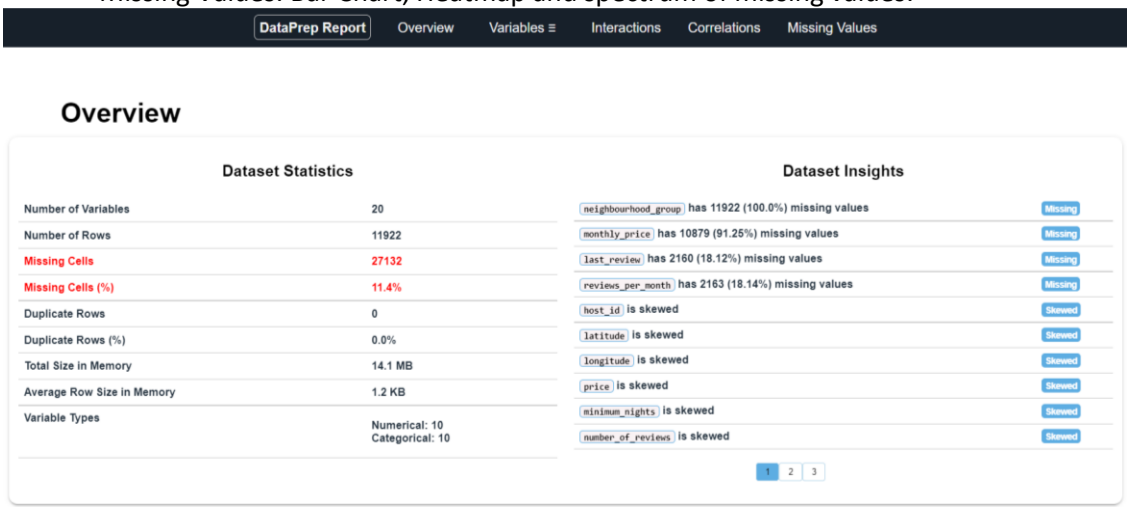
```
In [2]: df_Listings = pd.read_csv('listings.csv')
        df_Listings.head()
```

Out[2]:

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	city	latitude	longitude	property_type	room_type	
0	6	Large Craftsmen w/ yard ~ Perfect for families	29	Sara	NaN	North Hills	San Diego	32.753993	-117.129705	House	Entire home/apt	
1	5570	Ocean front condo on the sand	8435	Jef Karchin'S MISSIONBEACHRETREAT	NaN	Mission Bay	San Diego	32.784304	-117.252578	Condominium	Entire home/apt	TV
2	8095	Sunset Cliffs Studio	270	Marin	NaN	Ocean Beach	San Diego	32.735170	-117.243793	Guesthouse	Entire home/apt	
3	39516	Art Studio Retreat/Rustic Cabin	169649	Chris And Jean	NaN	North Hills	San Diego	32.731884	-117.119180	Tiny house	Entire home/apt	{Inte
4	45429	OB cottage SD--view on waterway	197919	Melissa	NaN	Loma Portal	San Diego	32.748768	-117.229371	House	Entire home/apt	{T

3.1.4 Exploratory Data Analysis (EDA) –

- "Exploratory Data Analysis" (EDA) is a "Data Exploration" step in the Data Analysis Process, where a number of techniques are used to better understand the dataset being used.
- Understanding the Dataset can refer to a number of things including but not limited to…
  - Extracting Important "Variables".
  - Identifying "Outliers", "Missing Values", or "Human Error".
  - Understanding the Relationships between variables.
  - Ultimately, maximizing our insights of a dataset and minimizing potential "Error" that may occur later in the process.
- In other words, it will gives you a better Understanding of the "Variables" and the "Relationships" between them.
- Here, we make use of dataprep module to automate our EDA process.
- It provides the following information:
  - Overview: detect the types of columns in a DataFrame.
  - Variables: variable type, unique values, distinct count, missing values
  - Quartile statistics like minimum value, Q1, median, Q3, maximum, range, interquartile range
  - Descriptive statistics like mean, mode, standard deviation, sum, median absolute deviation, coefficient of variation, kurtosis, skewness.
  - Correlations: highlighting of highly correlated variables, Spearman, Pearson and Kendall matrices
  - Missing Values: Bar Chart, Heatmap and spectrum of missing values.



3.1.5 Data Pre-processing, Data Cleaning & Imputation (Handling the Categorical & Numerical Variables) –

Data pre-processing is a process of preparing the raw data and making it suitable for our analysis purpose, where we have to do lot of Data Cleaning, handle the missing values by using appropriate imputation techniques and based on that variable nature i.e. either of Categorical & Numerical variable. Here, in this project, we have done the substitution/imputation of missing values using either mean, median or mode according to the nature of those variables. Moreover, we also removed the columns which are does not participate in our analysis.

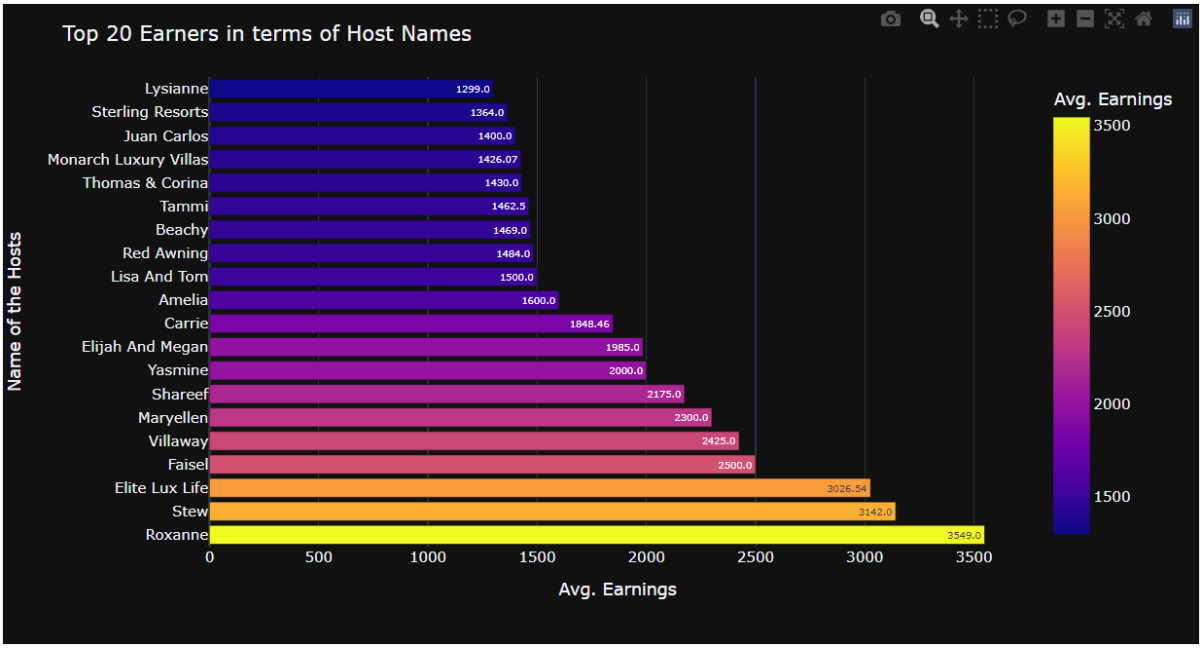
3.1.6 Analyse the Data –

Once the pre-processing is done, we are good to go with our actual analysis where we write lines of codes and logics to prepare our data as per the defined use cases.

3.1.7 Visualize & Share Meaningful Insights –

Finally, it's time to turn our data into some sort of visual representation. In short, Data visualization is the process of translating large data sets and metrics into charts, graphs and other visuals such as Bar Plot, Pie Chart, Heat map, Box Plot, Scatter Plot, and many more. The resulting visual representation of data makes it easier to identify and share insights about the information represented in the data.

Here is the beautiful glimpse of one of our visuals are –



All those different analysis help out to make better business decisions and help analyse customer trends and satisfaction, which can lead to new and better products and services.

4 Technology Stack

Data Manipulation & Mathematical Computation Library	Pandas, NumPy
Visualization Library	Matplotlib, Seaborn, Plotly, WordCloud, etc
EDA	dataprep
Dataset	.CSV Format
IDE	Jupyter Notebook