Finlatics
Programme: Business Analytics
Project 2

Suryanshu

February 16, 2024

Contents

Business Analytics

Project 2

1.1   Introduction

The problem statement requires us to do two basic things:

1. **Data Processing:** Convert the unstructured-map data, into structured data-excel sheet. And from this find various information regarding the data set like:

   - State with the hightest number of hotels
   - Filtering out the states with three types of climates
   - Average number of days the rainy season lasts in Indian states

2. **Analytics:** After processing the data, it must be interpretted and analysed. From this analysis we must suggest:

   - Among the northeastern states which are best to set up a hotel
   - Which is the best state for setting up a hotel

Here is the map provided.

Figure 1.1: Map of Climatic Regions in India

1.2 Data Processing

The first step here is to process the Hotel_Dataset provided. Then we convert the unstructured-map data into structure-excel sheet(the different climates and the onset and retreat of monsoon in each state).

1.2.1 Hotel_Dataset

	A	B
1	**States** ⌄	**Max of No of Hotels**
2	Kerala	672000
3	Tamil Nadu	116000
4	Gujarat	107000
5	Maharashtra	102000
6	West Bengal	87000
7	Himachal Pradesh	83000
8	Rajasthan	72000
9	Uttar Pradesh	58000
10	Uttarakhand	50000
11	Haryana	48000
12	Goa	44000
13	Karnataka	40000
14	Telangana	34000
15	Andhra Pradesh	34000
16	Chhattisgarh	25000
17	Assam	25000
18	Odisha	20000
19	Madhya Pradesh	14000
20	Sikkim	12000
21	Punjab	11000
22	Arunachal Pradesh	5000
23	Jharkhand	4000
24	Bihar	4000
25	Manipur	3000
26	Meghalaya	2000
27	Mizoram	1000
28	Tripura	1000
29	Nagaland	1000
30	**Grand Total**	**672000**

Figure 1.2: Hotel Dataset in a sorted pivot table

From this we learn that **Kerala** has the hightest number of hotels, i.e. 672000.

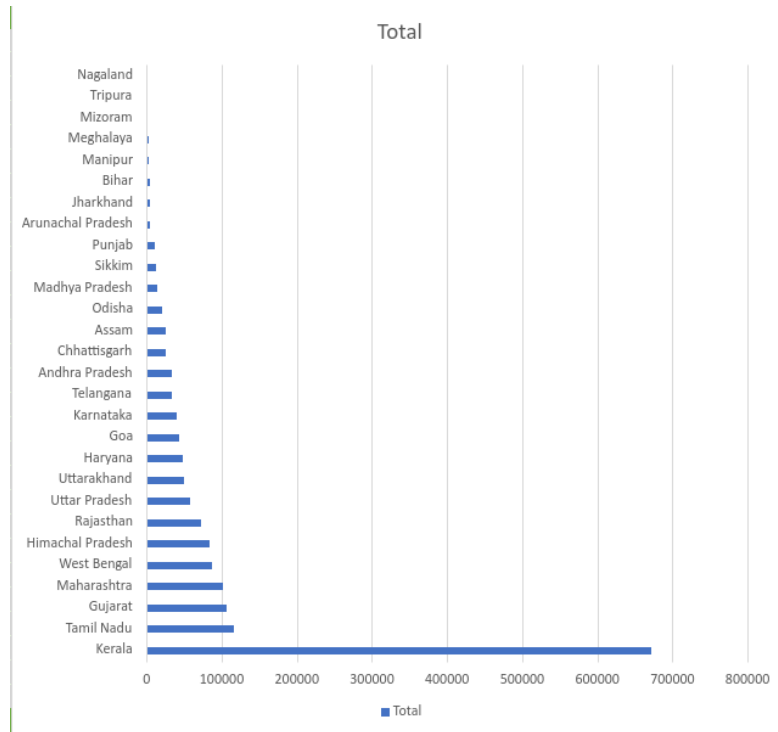This data can be presented via a bar graph which compares all the states.

Figure 1.3: Hotel Dataset in a bar graph

1.2.2 Climatic Conditions

Here we convert the map into two excelsheets- first for the types of climates in a state, second the number of days of rainy season.

	States	Sum of Tropical Monsoon	Sum of Tropical Savannah	Sum of Arid Steppe	Sum of Humid Subtropical	Sum of Montane Climate	Sum of Hot Deserts	Sum of Total	Total
1									
2	Andhra Pradesh	0	1	1	0	0	0	2	
3	Arunachal Pradesh	0	0	0	1	1	0	2	1
4	Assam	0	0	0	1	1	0	2	2
5	Bihar	0	0	0	1	0	0	1	3
6	Chhattisgarh	0	1	0	1	0	0	2	4
7	Goa	1	0	0	0	0	0	1	
8	Gujarat	0	0	1	0	0	1	2	
9	Haryana	0	0	1	0	0	1	2	
10	Himachal Pradesh	0	0	1	0	1	0	2	
11	Jharkhand	0	1	0	1	0	0	2	
12	Karnataka	1	1	0	0	0	0	2	
13	Kerala	1	1	1	0	0	0	3	
14	Madhya Pradesh	0	1	1	1	0	0	3	
15	Maharashtra	1	1	1	1	0	0	4	
16	Manipur	0	0	0	1	0	0	1	
17	Meghalaya	0	0	0	1	0	0	1	
18	Mizoram	0	0	0	1	0	0	1	
19	Nagaland	0	0	0	1	0	0	1	
20	Odisha	0	1	0	1	0	0	2	
21	Punjab	0	0	1	0	0	1	2	
22	Rajasthan	0	0	1	0	0	1	2	
23	Sikkim	0	0	0	0	1	0	1	
24	Tamil Nadu	1	1	1	0	0	0	3	
25	Telangana	0	1	1	0	0	0	2	
26	Tripura	0	0	0	1	0	0	1	
27	Uttar Pradesh	0	0	1	1	0	0	2	
28	Uttarakhand	0	0	0	1	1	0	2	
29	West Bengal	0	1	0	1	1	0	3	
30	Grand Total	5	11	12	16	6	4	54	

Figure 1.4: Weather Data into a tabular form

From this we filter out the states which have three types of climatic conditions.

	States	Sum of Tropical Monsoon	Sum of Tropical Savannah	Sum of Arid Steppe	Sum of Humid Subtropical	Sum of Montane Climate	Sum of Hot Deserts	Sum of Total	Total
1									
2	Kerala	1	1	1	0	0	0	3	1
3	Madhya Pradesh	0	1	1	1	0	0	3	2
4	Tamil Nadu	1	1	1	0	0	0	3	3
5	West Bengal	0	1	0	1	1	0	3	4
6	Grand Total	2	4	3	2	1	0	12	

Figure 1.5: Filtered Data

We find that 4 states namely-

- Kerala

- Madhya Pradesh

- Tamil Nadu

- West Bengal

have three types of climatic condtions.

Now we follow the datelines presented and convert the monsoon onset and retreat into structured data.

	A	B	C	D	E
1	Serial No.	States	Monsoon Onset [Earliest]	Monsoon Retreat [Latest]	Days
2	1	Andhra Pradesh	1-Jun	15-Oct	136
3	2	Arunachal Pradesh	5-Jun	1-Oct	118
4	3	Assam	5-Jun	1-Oct	118
5	4	Bihar	10-Jun	1-Oct	113
6	5	Chhattisgarh	5-Jun	15-Oct	132
7	6	Goa	10-Jun	1-Oct	113
8	7	Gujarat	10-Jun	1-Oct	113
9	8	Haryana	1-Jul	1-Oct	92
10	9	Himachal Pradesh	1-Jul	1-Oct	92
11	10	Jharkhand	10-Jun	1-Oct	113
12	11	Karnataka	1-Jun	15-Oct	136
13	12	Kerala	1-Jun	15-Oct	136
14	13	Madhya Pradesh	10-Jun	1-Oct	113
15	14	Maharashtra	5-Jun	1-Oct	118
16	15	Manipur	1-Jun	1-Oct	122
17	16	Meghalaya	1-Jun	1-Oct	122
18	17	Mizoram	1-Jun	1-Oct	122
19	18	Nagaland	5-Jun	1-Oct	118
20	19	Odisha	5-Jun	1-Oct	118
21	20	Punjab	1-Jul	15-Sep	76
22	21	Rajasthan	15-Jun	1-Oct	108
23	22	Sikkim	10-Jun	1-Oct	113
24	23	Tamil Nadu	1-Jun	15-Oct	136
25	24	Telangana	5-Jun	15-Oct	132
26	25	Tripura	1-Jun	1-Oct	122
27	26	Uttar Pradesh	15-Jun	1-Oct	108
28	27	Uttarakhand	1-Jul	15-Sep	76
29	28	West Bengal	5-Jun	1-Oct	118
30					115.5

Figure 1.6: Average number of days the rainy season lasts in Indian states

From this we calculate the average number of days the rainy season lasts in Indian states, which comes out to be 115.5 days.

1.2.3 Data on Northeastern States

In order to analyse the northeastern states of India, we'll separate their data from the rest and then reprocess it.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Northeastern States	Tropical Monsoon	Tropical Savannah	Arid Steppe	Humid Subtropical	Montane Climate	Hot Deserts	Total	Monsoon Onset	Monsoon Retreat	Days	No. of Hotels
2	Arunachal Pradesh	0	0	0	1	1	0	2	5-Jun	1-Oct	118	5000
3	Assam	0	0	0	1	1	0	2	5-Jun	1-Oct	118	25000
4	Manipur	0	0	0	1	0	0	1	1-Jun	1-Oct	122	3000
5	Meghalaya	0	0	0	1	0	0	1	1-Jun	1-Oct	122	2000
6	Mizoram	0	0	0	1	0	0	1	1-Jun	1-Oct	122	1000
7	Nagaland	0	0	0	1	0	0	1	5-Jun	1-Oct	118	1000
8	Sikkim	0	0	0	0	1	0	1	10-Jun	1-Oct	113	12000
9	Tripura	0	0	0	1	0	0	1	1-Jun	1-Oct	122	1000

Figure 1.7: Northeastern States' Data

We make a separate pivot table for the average of days and average number of hotels.

Northeastern States	Average of Days	Average of No. of Hotels
Arunachal Pradesh	118	5000
Assam	118	25000
Manipur	122	3000
Meghalaya	122	2000
Mizoram	122	1000
Nagaland	118	1000
Sikkim	113	12000
Tripura	122	1000
Grand Total	**119.375**	**6250**

Figure 1.8: Hotel and Monsoon NE Pivot Table

We make a clustered bar graph to compare all the NE states in terms of rainy days and number of hotels
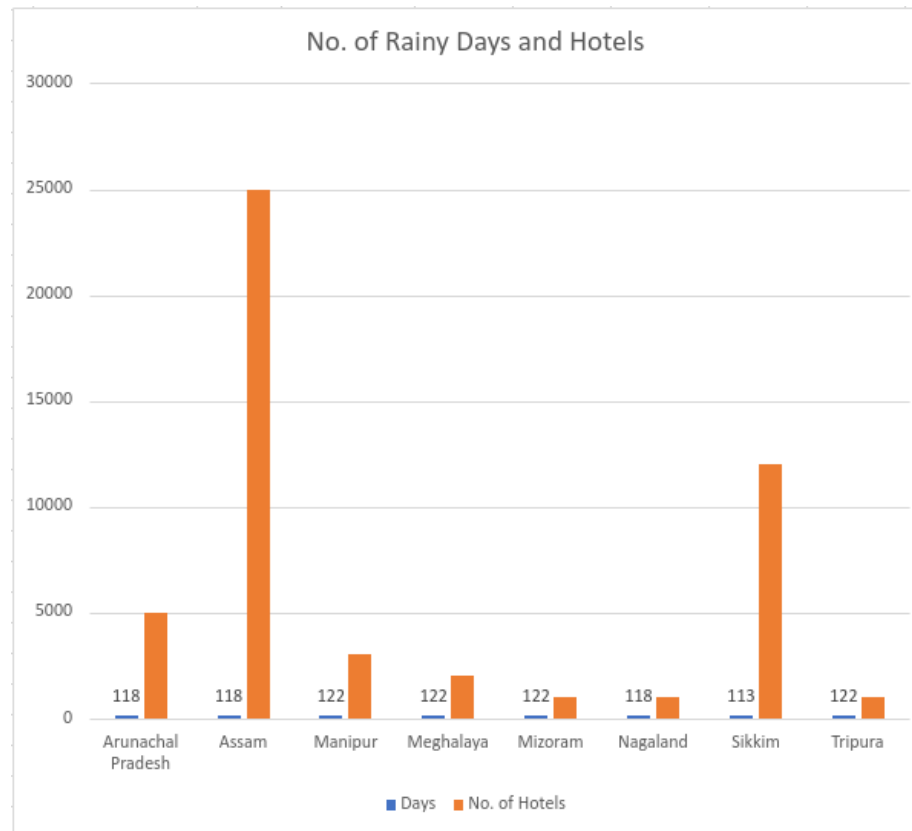

Figure 1.9: Number of Rainy Days and Hotels in NE states

We use a stacked bar graph to identify the proportion of different climatic conditions present in the NE region.
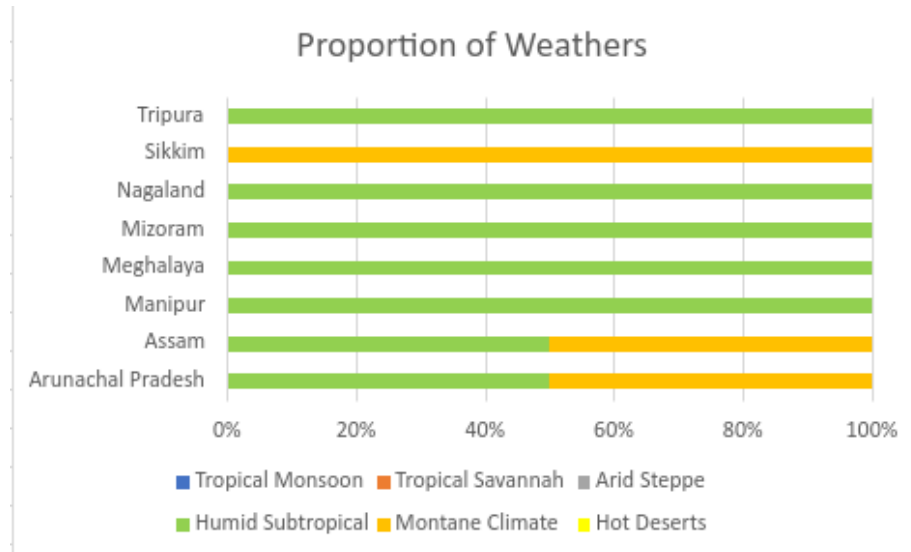

Figure 1.10: Proportion of Climatic Conditions in the NE region

This concludes our data processing part and we move on to analysing this data to draw conclusions out of it.

1.3 Analytics

We have three factors to analyse whether a state is optimal to set up a hotel or not.

- **Climatic Condtions:** We pick the state with moderate climate, if the the weather's too harsh, tourists are unlikely to show up.

- **Monsoon:** Tourists would rarely travel in monsoon, hence the state with fews rainy days would be preferred.

- **Number of already established hotels:** If the competition is too hgh that would negatively impact the business.

1.3.1 Best Northeastern state to set up a hotel

After analysing the data we find that only Sikkim(12000) and Assam(25000) have a huge number of hotels. Either we can establish hotels in those states and face fierce competition or we can invest in **Meghalaya**(2000) and **Arunachal Pradesh**(5000) with moderate competition and hence still an untapped market.

Even weather seems favourable for Arunachal Pradesh and Meghalaya which recieve rains for 118 and 122 days respectively with a humid subtropical climate.

1.3.2 Best state to set up a hotel

As a hotel requires a good influx of tourists in order to thrive, we must look at what attracts tourists in the first place. While a tourist attraction is a must, but there must be an already state established infrastructure to support tourism.

We find that **Karnataka** with all its natural beauty, tourist and religious spots still has on average fewer hotels (hence less competition) than say, Kerala or Tamil nadu. The weather is a mix of wet and dry with tropical monsoon and savannah with it rainy 136 days.