

Data Science Bootcamp T5

Summary :

In this project I chose a data set about an airline and my goal is to predict the level of passenger satisfaction.

Data set :

The data size is 3.04 MB

- ID
- Gender
- Customer
- Age
- Type of travel
- Class
- Flight Distance
- Inflight wifi service
- Departure/Arrival time convenient
- Ease of Online booking
- Gate location
- Food and drink
- Online boarding
- Seat comfort
- Inflight entertainment
- On-board service
- Leg room service
- Baggage handling
- Checkin service
- Inflight service
- Cleanliness
- Departure Delay in Minutes
- Arrival Delay in Minutes
- Target: Satisfaction

Work :

- Drop missing data

```
[10]: df.dropna()
```

- Removing Duplicated rows

```
[14]: df.duplicated()
```

```
[14]: 0      False
      1      False
      2      False
      3      False
      4      False
      ...
      71544 False
      71545 False
      71546 False
      71547 False
      71548 False
      Length: 71549, dtype: bool
```

Add co
tools

```
[15]: df.duplicated().sum()
```

```
[15]: 0
```

- Summary

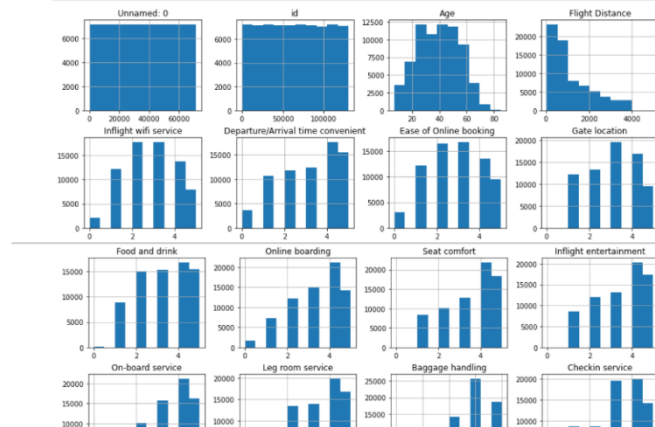
```
[16]: df.describe().T
```

```
[16]:
```

		count	mean	std	min	25%	50%	75%	max
Unnamed: 0	71549.0	35774.000000	20654.561542	0.0	17887.0	35774.0	53661.0	71548.0	
id	71549.0	64815.304840	37475.852677	2.0	32406.0	64714.0	97240.0	129880.0	
Age	71549.0	39.382857	15.099017	7.0	27.0	40.0	51.0	85.0	
Flight Distance	71549.0	1187.926722	996.140346	31.0	413.0	840.0	1739.0	4983.0	
Inflight wifi service	71549.0	2.733204	1.330131	0.0	2.0	3.0	4.0	5.0	
Departure/Arrival time convenient	71549.0	3.059735	1.527065	0.0	2.0	3.0	4.0	5.0	
Ease of Online booking	71549.0	2.756391	1.399490	0.0	2.0	3.0	4.0	5.0	
Gate location	71549.0	2.973934	1.278876	0.0	2.0	3.0	4.0	5.0	
Food and drink	71549.0	3.205585	1.331446	0.0	2.0	3.0	4.0	5.0	
Online boarding	71549.0	3.252359	1.349354	0.0	2.0	3.0	4.0	5.0	
Seat comfort	71549.0	3.442941	1.320747	0.0	2.0	4.0	5.0	5.0	
Inflight entertainment	71549.0	3.359027	1.333851	0.0	2.0	4.0	4.0	5.0	
On-board service	71549.0	3.381669	1.288191	0.0	2.0	4.0	4.0	5.0	
Leg room service	71549.0	3.348279	1.312775	0.0	2.0	4.0	4.0	5.0	
Baggage handling	71549.0	3.634069	1.180745	1.0	3.0	4.0	5.0	5.0	
Checkin service	71548.0	3.308338	1.263416	0.0	3.0	3.0	4.0	5.0	

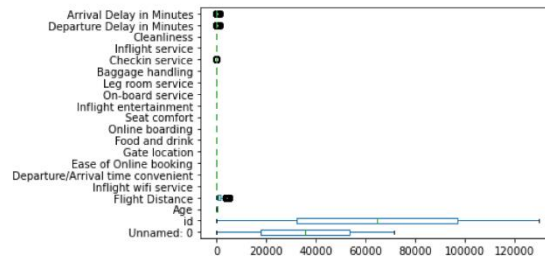
- Histogram

```
[19]: hist = df.hist(bins=10,figsize =(15,15))
```



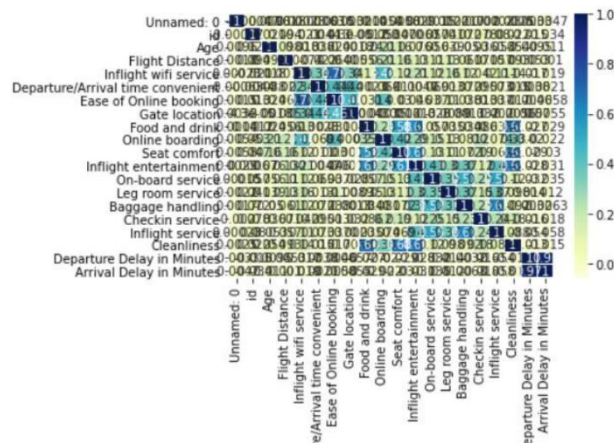
- Review Outliers
- Create box plot for all columns

```
[23]: bxp1t = df.boxplot(grid=False, vert=False, fontsize=10)
```

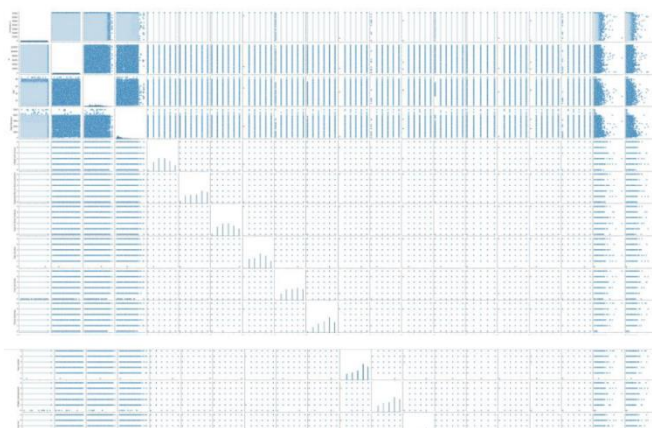


Identify outliers . the column at left side.

- Data Relation



- Pair plots , Evaluate the column distribution against each other columns



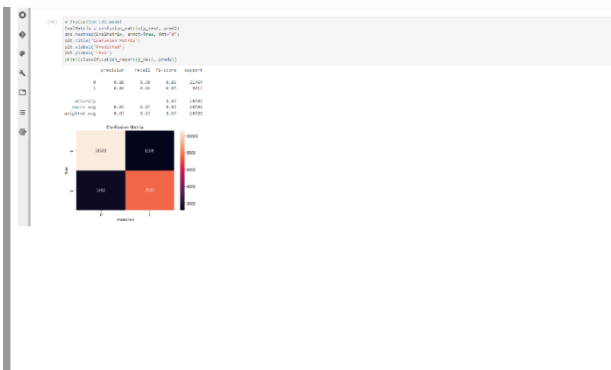
```

100 |> summarise(
101 |   # Create a new column 'Age' based on 'Age' and 'Sex'
102 |   Age = ifelse(Sex == "Male", Age + 1, Age),
103 |   # Create a new column 'Status' based on 'Age' and 'Sex'
104 |   Status = ifelse(Age > 18, "Adult", "Minor"),
105 |   # Create a new column 'Category' based on 'Age' and 'Sex'
106 |   Category = ifelse(Age > 18 & Sex == "Male", "A", "B"),
107 |   # Create a new column 'Value' based on 'Age' and 'Sex'
108 |   Value = ifelse(Age > 18 & Sex == "Male", 100, 50),
109 |   # Create a new column 'Label' based on 'Age' and 'Sex'
110 |   Label = ifelse(Age > 18 & Sex == "Male", "High", "Low")
111 | )
112 |> # Print the result
113 |> print(df)
114 |> # Print the column names
115 |> print(colnames(df))
116 |> # Print the data types of the columns
117 |> print(class(df))
118 |> # Print the dimensions of the data frame
119 |> print(dim(df))
120 |> # Print the first few rows of the data frame
121 |> print(head(df))
122 |> # Print the last few rows of the data frame
123 |> print(tail(df))
124 |> # Print the unique values of a column
125 |> print(unique(df$Age))
126 |> # Print the frequency of values in a column
127 |> print(table(df$Age))
128 |> # Print the summary statistics of a column
129 |> print(summary(df$Age))
130 |> # Print the summary statistics of all columns
131 |> print(summary(df))
132 |> # Print the correlation matrix
133 |> print(cor(df))
134 |> # Print the scatter plot of two columns
135 |> plot(df$Age, df$Value)
136 |> # Print the box plot of a column
137 |> boxplot(df$Value)
138 |> # Print the histogram of a column
139 |> hist(df$Value)
140 |> # Print the mean of a column
141 |> print(mean(df$Value))
142 |> # Print the median of a column
143 |> print(median(df$Value))
144 |> # Print the mode of a column
145 |> print(mode(df$Value))
146 |> # Print the standard deviation of a column
147 |> print(sd(df$Value))
148 |> # Print the variance of a column
149 |> print(var(df$Value))
150 |> # Print the range of a column
151 |> print(range(df$Value))
152 |> # Print the minimum of a column
153 |> print(min(df$Value))
154 |> # Print the maximum of a column
155 |> print(max(df$Value))
156 |> # Print the quantiles of a column
157 |> print(quantiles(df$Value))
158 |> # Print the percentiles of a column
159 |> print(percentiles(df$Value))
160 |> # Print the deciles of a column
161 |> print(deciles(df$Value))
162 |> # Print the quartiles of a column
163 |> print(quartiles(df$Value))
164 |> # Print the skewness of a column
165 |> print(skewness(df$Value))
166 |> # Print the kurtosis of a column
167 |> print(kurtosis(df$Value))
168 |> # Print the moments of a column
169 |> print(moments(df$Value))
170 |> # Print the moments of all columns
171 |> print(moments(df))
172 |> # Print the moments of the first few columns
173 |> print(moments(df[, 1:3]))
174 |> # Print the moments of the last few columns
175 |> print(moments(df[, 4:6]))
176 |> # Print the moments of a subset of columns
177 |> print(moments(df[, c("Age", "Status")]))
178 |> # Print the moments of a specific column
179 |> print(moments(df$Age))
180 |> # Print the moments of a specific row
181 |> print(moments(df[1,]))
182 |> # Print the moments of a specific column and row
183 |> print(moments(df$Age, df[1,]))
184 |> # Print the moments of a specific column and row subset
185 |> print(moments(df$Age, df[1:5,]))
186 |> # Print the moments of a specific column and row subset
187 |> print(moments(df$Age, df[1:5, 1:3]))
188 |> # Print the moments of a specific column and row subset
189 |> print(moments(df$Age, df[1:5, 1:3, 1:3]))
190 |> # Print the moments of a specific column and row subset
191 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3]))
192 |> # Print the moments of a specific column and row subset
193 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3]))
194 |> # Print the moments of a specific column and row subset
195 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3]))
196 |> # Print the moments of a specific column and row subset
197 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
198 |> # Print the moments of a specific column and row subset
199 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
200 |> # Print the moments of a specific column and row subset
201 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
202 |> # Print the moments of a specific column and row subset
203 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
204 |> # Print the moments of a specific column and row subset
205 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
206 |> # Print the moments of a specific column and row subset
207 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
208 |> # Print the moments of a specific column and row subset
209 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
210 |> # Print the moments of a specific column and row subset
211 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
212 |> # Print the moments of a specific column and row subset
213 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
214 |> # Print the moments of a specific column and row subset
215 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
216 |> # Print the moments of a specific column and row subset
217 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
218 |> # Print the moments of a specific column and row subset
219 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
220 |> # Print the moments of a specific column and row subset
221 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
222 |> # Print the moments of a specific column and row subset
223 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
224 |> # Print the moments of a specific column and row subset
225 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
226 |> # Print the moments of a specific column and row subset
227 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
228 |> # Print the moments of a specific column and row subset
229 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
230 |> # Print the moments of a specific column and row subset
231 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
232 |> # Print the moments of a specific column and row subset
233 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
234 |> # Print the moments of a specific column and row subset
235 |> print(moments(df$Age, df[1:5, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3, 1:3]))
236 |> # Print the moments of a specific column and row subset
23
```

[illegible]

```

100 # 30 days
101
102 # Create X
103 X = pd.DataFrame()
104
105 # Add features
106 X['Month'] = 0
107 X['Day'] = 0
108 X['Hour'] = 0
109 X['Temperature'] = 0
110 X['Humidity'] = 0
111 X['WindSpeed'] = 0
112 X['CloudCover'] = 0
113 X['Sunshine'] = 0
114 X['Rainfall'] = 0
115 X['Snowfall'] = 0
116 X['Icefall'] = 0
117 X['Fogfall'] = 0
118 X['Thunderfall'] = 0
119 X['Hailfall'] = 0
120 X['Sleetfall'] = 0
121 X['Snowfall'] = 0
122 X['Icefall'] = 0
123 X['Fogfall'] = 0
124 X['Thunderfall'] = 0
125 X['Hailfall'] = 0
126 X['Sleetfall'] = 0
127 X['Snowfall'] = 0
128 X['Icefall'] = 0
129 X['Fogfall'] = 0
130 X['Thunderfall'] = 0
131 X['Hailfall'] = 0
132 X['Sleetfall'] = 0
133 X['Snowfall'] = 0
134 X['Icefall'] = 0
135 X['Fogfall'] = 0
136 X['Thunderfall'] = 0
137 X['Hailfall'] = 0
138 X['Sleetfall'] = 0
139 X['Snowfall'] = 0
140 X['Icefall'] = 0
141 X['Fogfall'] = 0
142 X['Thunderfall'] = 0
143 X['Hailfall'] = 0
144 X['Sleetfall'] = 0
145 X['Snowfall'] = 0
146 X['Icefall'] = 0
147 X['Fogfall'] = 0
148 X['Thunderfall'] = 0
149 X['Hailfall'] = 0
150 X['Sleetfall'] = 0
151 X['Snowfall'] = 0
152 X['Icefall'] = 0
153 X['Fogfall'] = 0
154 X['Thunderfall'] = 0
155 X['Hailfall'] = 0
156 X['Sleetfall'] = 0
157 X['Snowfall'] = 0
158 X['Icefall'] = 0
159 X['Fogfall'] = 0
160 X['Thunderfall'] = 0
161 X['Hailfall'] = 0
162 X['Sleetfall'] = 0
163 X['Snowfall'] = 0
164 X['Icefall'] = 0
165 X['Fogfall'] = 0
166 X['Thunderfall'] = 0
167 X['Hailfall'] = 0
168 X['Sleetfall'] = 0
169 X['Snowfall'] = 0
170 X['Icefall'] = 0
171 X['Fogfall'] = 0
172 X['Thunderfall'] = 0
173 X['Hailfall'] = 0
174 X['Sleetfall'] = 0
175 X['Snowfall'] = 0
176 X['Icefall'] = 0
177 X['Fogfall'] = 0
178 X['Thunderfall'] = 0
179 X['Hailfall'] = 0
180 X['Sleetfall'] = 0
181 X['Snowfall'] = 0
182 X['Icefall'] = 0
183 X['Fogfall'] = 0
184 X['Thunderfall'] = 0
185 X['Hailfall'] = 0
186 X['Sleetfall'] = 0
187 X['Snowfall'] = 0
188 X['Icefall'] = 0
189 X['Fogfall'] = 0
190 X['Thunderfall'] = 0
191 X['Hailfall'] = 0
192 X['Sleetfall'] = 0
193 X['Snowfall'] = 0
194 X['Icefall'] = 0
195 X['Fogfall'] = 0
196 X['Thunderfall'] = 0
197 X['Hailfall'] = 0
198 X['Sleetfall'] = 0
199 X['Snowfall'] = 0
200 X['Icefall'] = 0
201 X['Fogfall'] = 0
202 X['Thunderfall'] = 0
203 X['Hailfall'] = 0
204 X['Sleetfall'] = 0
205 X['Snowfall'] = 0
206 X['Icefall'] = 0
207 X['Fogfall'] = 0
208 X['Thunderfall'] = 0
209 X['Hailfall'] = 0
210 X['Sleetfall'] = 0
211 X['Snowfall'] = 0
212 X['Icefall'] = 0
213 X['Fogfall'] = 0
214 X['Thunderfall'] = 0
215 X['Hailfall'] = 0
216 X['Sleetfall'] = 0
217 X['Snowfall'] = 0
218 X['Icefall'] = 0
219 X['Fogfall'] = 0
220 X['Thunderfall'] = 0
221 X['Hailfall'] = 0
222 X['Sleetfall'] = 0
223 X['Snowfall'] = 0
224 X['Icefall'] = 0
225 X['Fogfall'] = 0
226 X['Thunderfall'] = 0
227 X['Hailfall'] = 0
228 X['Sleetfall'] = 0
229 X['Snowfall'] = 0
230 X['Icefall'] = 0
231 X['Fogfall'] = 0
232 X['Thunderfall'] = 0
233 X['Hailfall'] = 0
234 X['Sleetfall'] = 0
235 X['Snowfall'] = 0
236 X['Icefall'] = 0
237 X['Fogfall'] = 0
238 X['Thunderfall'] = 0
239 X['Hailfall'] = 0
240 X['Sleetfall'] = 0
241 X['Snowfall'] = 0
242 X['Icefall'] = 0
243 X['Fogfall'] = 0
244 X['Thunderfall'] = 0
245 X['Hailfall'] = 0
246 X['Sleetfall'] = 0
247 X['Snowfall'] = 0
248 X['Icefall'] = 0
249 X['Fogfall'] = 0
250 X['Thunderfall'] = 0
251 X['Hailfall'] = 0
252 X['Sleetfall'] = 0
253 X['Snowfall'] = 0
254 X['Icefall'] = 0
255 X['Fogfall'] = 0
256 X['Thunderfall'] = 0
257 X['Hailfall'] = 0
258 X['Sleetfall'] = 0
259 X['Snowfall'] = 0
260 X['Icefall'] = 0
261 X['Fogfall'] = 0
262 X['Thunderfall'] = 0
263 X['Hailfall'] = 0
264 X['Sleetfall'] = 0
265 X['Snowfall'] = 0
266 X['Icefall'] = 0
267 X['Fogfall'] = 0
268 X['Thunderfall'] = 0
269 X['Hailfall'] = 0
270 X['Sleetfall'] = 0
271 X['Snowfall'] = 0
272 X['Icefall'] = 0
273 X['Fogfall'] = 0
274 X['Thunderfall'] = 0
275 X['Hailfall'] = 0
276 X['Sleetfall'] = 0
277 X['Snowfall'] = 0
278 X['Icefall'] = 0
279 X['Fogfall'] = 0
280 X['Thunderfall'] = 0
281 X['Hailfall'] = 0
282 X['Sleetfall'] = 0
283 X['Snowfall'] = 0
284 X['Icefall'] = 0
285 X['Fogfall'] = 0
286 X['Thunderfall'] = 0
287 X['Hailfall'] = 0
288 X['Sleetfall'] = 0
289 X['Snowfall'] = 0
290 X['Icefall'] = 0
291 X['Fogfall'] = 0
292 X['Thunderfall'] = 0
293 X['Hailfall'] = 0
294 X['Sleetfall'] = 0
295 X['Snowfall'] = 0
296 X['Icefall'] = 0
297 X['Fogfall'] = 0
298 X['Thunderfall'] = 0
299 X['Hailfall'] = 0
300 X['Sleetfall'] = 0
301 X['Snowfall'] = 0
302 X['Icefall'] = 0
303 X['Fogfall'] = 0
304 X['Thunderfall'] = 0
305 X['Hailfall'] = 0
306 X['Sleetfall'] = 0
307 X['Snowfall'] = 0
308 X['Icefall'] = 0
309 X['Fogfall'] = 0
310 X['Thunderfall'] = 0
311 X['Hailfall'] = 0
312 X['Sleetfall'] = 0
313 X['Snowfall'] = 0
314 X['Icefall'] = 0
315 X['Fogfall'] = 0
316 X['Thunderfall'] = 0
317 X['Hailfall'] = 0
318 X['Sleetfall'] = 0
319 X['Snowfall'] = 0
320 X['Icefall'] = 0
321 X['Fogfall'] = 0
322 X['Thunderfall'] = 0
323 X['Hailfall'] = 0
324 X['Sleetfall'] = 0
325 X['Snowfall'] = 0
326 X['Icefall'] = 0
327 X['Fogfall'] = 0
328 X['Thunderfall'] = 0
329 X['Hailfall'] = 0
330 X['Sleetfall'] = 0
331 X['Snowfall'] = 0
332 X['Icefall'] = 0
333 X['Fogfall'] = 0
334 X['Thunderfall'] = 0
335 X['Hailfall'] = 0
336 X['Sleetfall'] = 0
337 X['Snowfall'] = 0
338 X['Icefall'] = 0
339 X['Fogfall'] = 0
340 X['Thunderfall'] = 0
341 X['Hailfall'] = 0
342 X['Sleetfall'] = 0
343 X['Snowfall'] = 0
344 X['Icefall'] = 0
345 X['Fogfall'] = 0
346 X['Thunderfall'] = 0
347 X['Hailfall'] = 0
348 X['Sleetfall'] = 0
349 X['Snowfall'] = 0
350 X['Icefall'] = 0
351 X['Fogfall'] = 0
352 X['Thunderfall'] = 0
353 X['Hailfall'] = 0
354 X['Sleetfall'] = 0
355 X['Snowfall'] = 0
356 X['Icefall'] = 0
357 X['Fogfall'] = 0
358 X['Thunderfall'] = 0
359 X['Hailfall'] = 0
360 X['Sleetfall'] = 0
361 X['Snowfall'] = 0
362 X['Icefall'] = 0
363 X['Fogfall'] = 0
364 X['Thunderfall'] = 0
365 X['Hailfall'] = 0
366 X['Sleetfall'] = 0
367 X['Snowfall'] = 0
368 X['Icefall'] = 0
369 X['Fogfall'] = 0
370 X['Thunderfall'] = 0
371 X['Hailfall'] = 0
372 X['Sleetfall'] = 0
373 X['Snowfall'] = 0
374 X['Icefall'] = 0
375 X['Fogfall'] = 0
376 X['Thunderfall'] = 0
377 X['Hailfall'] = 0
378 X['Sleetfall'] = 0
379 X['Snowfall'] = 0
380 X['Icefall'] = 0
381 X['Fogfall'] = 0
382 X['Thunderfall'] = 0
383 X['Hailfall'] = 0
384 X['Sleetfall'] = 0
385 X['Snowfall'] = 0
386 X['Icefall'] = 0
387 X['Fogfall'] = 0
388 X['Thunderfall'] = 0
389 X['Hailfall'] = 0
390 X['Sleetfall'] = 0
391 X['Snowfall'] = 0
392 X['Icefall'] = 0
393 X['Fogfall'] = 0
394 X['Thunderfall'] = 0
395 X['Hailfall'] = 0
396 X['Sleetfall'] = 0
397 X['Snowfall'] = 0
398 X['Icefall'] = 0
399 X['Fogfall'] = 0
400 X['Thunderfall'] = 0
401 X['Hailfall'] = 0
402 X['Sleetfall'] = 0
403 X['Snowfall'] = 0
404 X['Icefall'] = 0
405 X['Fogfall'] = 0
406 X['Thunderfall'] = 0
407 X['Hailfall'] = 0
408 X['Sleetfall'] = 0
409 X['Snowfall'] = 0
410 X['Icefall'] = 0
411 X['Fogfall'] = 0
412 X['Thunderfall'] = 0
413 X['Hailfall'] = 0
414 X['Sleetfall'] = 0
415 X['Snowfall'] = 0
416 X['Icefall'] = 0
417 X['Fogfall'] = 0
418 X['Thunderfall'] = 0
419 X['Hailfall'] = 0
420 X['Sleetfall'] = 0
421 X['Snowfall'] = 0
422 X['Icefall'] = 0
423 X['Fogfall'] = 0
424 X['Thunderfall'] = 0
425 X['Hailfall'] = 0
426 X['Sleetfall'] = 0
427 X['Snowfall'] = 0
428 X['Icefall'] = 
```



Conclusion :

During the four weeks of learning, I applied what I learned in my project of data cleaning, analysis and modeling to suit my output based on my data set.