```
[In]: import pandas as pd
[In]: import numpy as np
[In]: pd.set_option('display.max_columns', 30)
[In]: pd.set_option('display.max_rows', 30)
[In]: %matplotlib inline
[In]: %pylab inline
[In]: %matplotlib inline
[In]: pd.pd.read_csv('C:\Space_Corrected.csv')
```

```
[Out]:
   Unnamed: 0.1  Unnamed: 0  Company Name                 Location                 Datum                  Detail  Status Rocket  Rocket  Status Mission
0             0           0         Arianespace               Kourou  Fri Aug 07, 2020 05:12 UTC   Falcon 9 Block 5 | Starlink V1 L9 & BlueOrig  Staus Active  811.0  Success
1             1           1               CASC   Site 9401 (SLS-2), Jiuquan Satellite Launch Ce...  Thu Aug 06, 2020 04:01 UTC                 Long March 2D | Gaofen-9 04 & Q-SAT  Status Active   29.75  Success
2             2           2               SpaceX                 Pad A, Boca Chica, Texas, USA  Thu Aug 06, 2020 07:37 UTC                      Starship Prototype | 150 Meter Hop  Status Active     NaN  Failure
3             3           3         Roscosmos   Site 200/39, Baikonur Cosmodrome, Kazakhstan  Thu Jul 30, 2020 21:29 UTC  Proton-M/Briz-M | Progress MS & Progress ISS  Status Active   415.0  Success
4             4           4               ULA                 SLC-41, Cape Canaveral AFS, Florida, USA  Thu Jul 30, 2020 11:50 UTC                                Atlas V 541 | Perseverance  Status Active   145.0  Success

...           ...         ...                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                

4224  4224  4224  US Navy  LC-18A, Cape Canaveral AFS, Florida, USA  Wed Feb 05, 1958 07:33 UTC  Vanguard | Vanguard TV3BU  Status Retired  NaN  Failure
4225  4225  4225  AMBA  LC-26A, Cape Canaveral AFS, Florida, USA  Sat Feb 01, 1958 03:48 UTC  Juno I | Explorer 1  Status Retired  NaN  Failure
4226  4226  4226  US Navy  LC-18A, Cape Canaveral AFS, Florida, USA  Fri Dec 06, 1957 16:44 UTC  Vanguard | Vanguard TV3  Status Retired  NaN  Failure
4227  4227  4227  RVSN USSR  Site 31, Baikonur Cosmodrome, Kazakhstan  Sun Dec 01, 1957 02:30 UTC  Sputnik 8K71PS | Sputnik 2  Status Retired  NaN  Success
4228  4228  4228  RVSN USSR  Site 31, Baikonur Cosmodrome, Kazakhstan  Fri Oct 04, 1957 19:28 UTC  Sputnik 8K71PS | Sputnik 1  Status Retired  NaN  Success

4229 rows × 9 columns
```

```
[In]: df.drop(['Unnamed: 0', 'Unnamed: 0.1'], axis = 1)
[In]: df.head()
```

	Company Name	Location	Datum	Detail	Status Rocket	Rocket	Status Mission
0	Arianespace	Kourou	Fri Aug 07, 2020 05:12 UTC	Falcon 9 Block 5	Starlink V1 L9 & BlueOrig	Staus Active	811.0	Success
1	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Ce...	Thu Aug 06, 2020 04:01 UTC	Long March 2D	Gaofen-9 04 & Q-SAT	Status Active	29.75	Success
2	SpaceX	Pad A, Boca Chica, Texas, USA	Thu Aug 06, 2020 07:37 UTC	Starship Prototype	150 Meter Hop	Status Active	NaN	Failure
3	Roscosmos	Site 200/39, Baikonur Cosmodrome, Kazakhstan	Thu Jul 30, 2020 21:29 UTC	Proton-M/Briz-M	Progress MS & Progress ISS	Status Active	415.0	Success
4	ULA	SLC-41, Cape Canaveral AFS, Florida, USA	Thu Jul 30, 2020 11:50 UTC	Atlas V 541	Perseverance	Status Active	145.0	Success

```
[In]: df.describe()
```

	Company Name	Location	Datum	Detail	Status Rocket	Rocket	Status Mission
count	4229	4229	4229	4229	4229	964	4229
unique	56	137	4200	1276	2	58	4
top	RVSN USSR	Site 31, Baikonur Cosmodrome, Kazakhstan	Mon Nov 05, 2008 00:16 UTC	Cosmos 3M(8K66)	SKR & Shuttle	StatusRetired	450.0	Success
freq	177	158	2	5	3534	298	3879

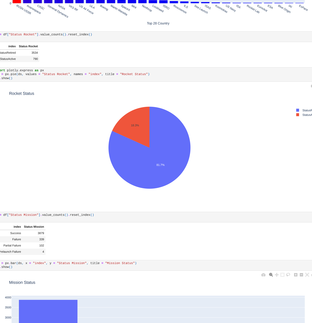
```
[In]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4229 entries, 0 to 4229
Data columns (total 9 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Company Name  4229 non-null   object
 1   Location      4229 non-null   object
 2   Datum         4229 non-null   object
 3   Detail        4229 non-null   object
 4   Status Rocket 4229 non-null   object
 5   Rocket        964 non-null    object
 6   Status Mission 4229 non-null  object
dtypes: float64(1), int64(2), object(6)
memory usage: 230.4 KB
```

```
[In]: df["Company Name"].value_counts().reset_index()[:20]
```

	Company Name
0	RVSN USSR	177
1	Arianespace	278
2	CASC	251
3	General Dynamics	251
4	NASA	203
5	US Navy	200
6	US Air Force	201
7	ULA	192
8	Boeing	138
9	Martin Marietta	114
10	SpaceX	110
11	MHI	84
12	Northrop	81
13	Lockheed	79
14	ISRO	76
15	Roscosmos	66
16	Rocketlab	61
17	Sea Launch	36
18	EADS	30
19	Kosmotras	22
20	US Navy	17
21	ISA	16
22	Pocket cat	13
23	Eurokot	13
24	Arianespace	13
25	Blue Origin	12
26	ULA	12
27	EuMas	10

```
[In]: import plotly.express as px
fig = px.bar(df, x="Company Name",
             color="Status Rocket",
             title="Number of Launches by Every Company",
             barmode="stack",
             height=500,
             width=1000,
             color_discrete_sequence=px.colors.qualitative.Pastel)
fig.show()
```



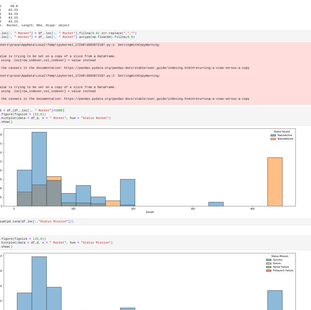

```
[In]: df.iloc[:,[0,1,2,3,4,5,6]].reset_index()
```

	Company Name	Location	Datum	Detail	Status Rocket	Rocket	Status Mission
0	Arianespace	Kourou	Fri Aug 07, 2020 05:12 UTC	Falcon 9 Block 5	Starlink V1 L9 & BlueOrig	StatusActive	811.0	Success
1	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Ce...	Thu Aug 06, 2020 04:01 UTC	Long March 2D	Gaofen-9 04 & Q-SAT	StatusActive	29.75	Success
2	SpaceX	Pad A, Boca Chica, Texas, USA	Thu Aug 06, 2020 07:37 UTC	Starship Prototype	150 Meter Hop	StatusActive	NaN	Failure
3	Roscosmos	Site 200/39, Baikonur Cosmodrome, Kazakhstan	Thu Jul 30, 2020 21:29 UTC	Proton-M/Briz-M	Progress MS & Progress ISS	StatusActive	415.0	Success
4	ULA	SLC-41, Cape Canaveral AFS, Florida, USA	Thu Jul 30, 2020 11:50 UTC	Atlas V 541	Perseverance	StatusActive	145.0	Success

```
[In]: df["Status Rocket"].value_counts().reset_index()
```

	Status Rocket
0	StatusRetired	3534
1	StatusActive	695

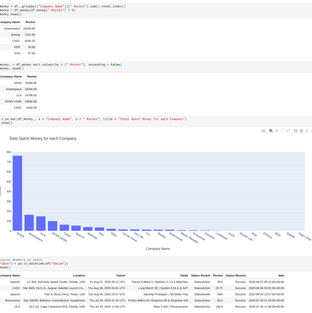
```
[In]: import plotly.express as px
fig = px.pie(df, values = "Status Rocket", names = "index", title = "Rocket Status")
fig.show()
```




```
[In]: df["Status Mission"].value_counts().reset_index()
```

	Status Mission
0	Success	3879
1	Failure	339
2	Partial Failure	10
3	Prelaunch Failure	1

```
[In]: fig = px.bar(df, x = "Status Mission", title = "Mission Status")
fig.show()
```




```
[In]: np.sum(df.iloc[df.iloc[:,7 'Rocket']])
```

```
[Out]:
0    1000
dtype: int64
```

```
[In]: df = df.dropna(subset=['Rocket'], axis = 'rows')
[In]: df.shape
```

```
[Out]:
964
```

```
[In]: df.iloc[:,[5 'Rocket']]
```

```
[Out]:
     Rocket
0     39.10
1     29.75
2    150.0
4    145.0
5     31.00
...     ...
4221  52.0
4222  52.0
4223  52.0
4224  0.0
[964 rows x 1 columns]: DeK, dtype: object
```

```
[In]: df.loc[:, 'Rocket'] = df.loc[:, 'Rocket'].str.replace(',', '')
df.loc[:, 'Rocket'] = df.loc[:, 'Rocket'].astype(np.float64).fillna(0.0)
```

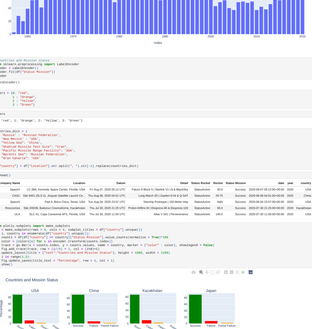
```
C:\Users\Yaroslav\AppData\Local\Temp\ipykernel_17240\2\3\8\8\3\2\7.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

C:\Users\Yaroslav\AppData\Local\Temp\ipykernel_17240\2\3\8\8\3\2\7.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

```
[In]: df.iloc[:,[5 'Rocket']]
[Out]: Figure(figsize = (15,6))
ax0.hist(df.iloc[:,[5 'Rocket']])
[Out]: [array([ 0.,  0.,  0.,  0.,  0.,  0.,  0.,  0.,  0.,  0.])]
```

```
[In]: fig = px.bar(df, x = "Status Mission")
fig.show()
```

```
[Out]:
0
```

```
[In]: figure(figsize = (15,6))
ax0.hist(df.iloc[:,[5 'Rocket']])
[Out]: [array([ 0.,  0.,  0.,  0.,  0.,  0.,  0.,  0.,  0.,  0.])]
```




```
[In]: # Total Spent Money for each company
df[['Rocket']]
```

	Company Name	Location	Datum	Detail	Status Rocket	Rocket	Status Mission
0	Arianespace	Kourou	Fri Aug 07, 2020 05:12 UTC	Falcon 9 Block 5	Starlink V1 L9 & BlueOrig	StatusActive	811.0	Success
1	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Ce...	Thu Aug 06, 2020 04:01 UTC	Long March 2D	Gaofen-9 04 & Q-SAT	StatusActive	29.75	Success
2	SpaceX	Pad A, Boca Chica, Texas, USA	Thu Aug 06, 2020 07:37 UTC	Starship Prototype	150 Meter Hop	StatusActive	NaN	Failure
3	Roscosmos	Site 200/39, Baikonur Cosmodrome, Kazakhstan	Thu Jul 30, 2020 21:29 UTC	Proton-M/Briz-M	Progress MS & Progress ISS	StatusActive	415.0	Success
4	ULA	SLC-41, Cape Canaveral AFS, Florida, USA	Thu Jul 30, 2020 11:50 UTC	Atlas V 541	Perseverance	StatusActive	145.0	Success

```
[In]: df_money = df[['Company Name']].groupby(['Rocket'].sum().reset_index()
df_money.sort_values(by = ['Rocket'], ascending = False)
[In]: df_money.head()
```

	Company Name	Rocket
58	NASA	9600.00
9	Arianespace	15240.00
25	ULA	14790.00
26	RVSN USSR	10800.00
7	CASC	6345.00

```
[In]: fig = px.bar(df_money, x = "Company Name", y = "Rocket", title = "Total Spent Money for each Company")
fig.show()
```



```
[In]: # Mission Number by year
df['year'] = df.iloc[:,[0]].astype('float')
df.head()
```

	Company Name	Location	Datum	Detail	Status Rocket	Rocket	Status Mission	date
0	Arianespace	Kourou	Fri Aug 07, 2020 05:12 UTC	Falcon 9 Block 5	Starlink V1 L9 & BlueOrig	StatusActive	811.0	Success	2020-08-07 05:12:00+00:00
1	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Ce...	Thu Aug 06, 2020 04:01 UTC	Long March 2D	Gaofen-9 04 & Q-SAT	StatusActive	29.75	Success	2020-08-06 04:01:00+00:00
2	SpaceX	Pad A, Boca Chica, Texas, USA	Thu Aug 06, 2020 07:37 UTC	Starship Prototype	150 Meter Hop	StatusActive	NaN	Failure	2020-08-06 07:37:00+00:00
3	Roscosmos	Site 200/39, Baikonur Cosmodrome, Kazakhstan	Thu Jul 30, 2020 21:29 UTC	Proton-M/Briz-M	Progress MS & Progress ISS	StatusActive	415.0	Success	2020-07-30 21:29:00+00:00
4	ULA	SLC-41, Cape Canaveral AFS, Florida, USA	Thu Jul 30, 2020 11:50 UTC	Atlas V 541	Perseverance	StatusActive	145.0	Success	2020-07-30 11:50:00+00:00

```
[In]: df['year'].apply(lambda datetime: datetime.year)
```

	Company Name	Location	Datum	Detail	Status Rocket	Rocket	Status Mission	date	year
0	Arianespace	Kourou	Fri Aug 07, 2020 05:12 UTC	Falcon 9 Block 5	Starlink V1 L9 & BlueOrig	StatusActive	811.0	Success	2020-08-07 05:12:00+00:00	2020
1	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Ce...	Thu Aug 06, 2020 04:01 UTC	Long March 2D	Gaofen-9 04 & Q-SAT	StatusActive	29.75	Success	2020-08-06 04:01:00+00:00	2020
2	SpaceX	Pad A, Boca Chica, Texas, USA	Thu Aug 06, 2020 07:37 UTC	Starship Prototype	150 Meter Hop	StatusActive	NaN	Failure	2020-08-06 07:37:00+00:00	2020
3	Roscosmos	Site 200/39, Baikonur Cosmodrome, Kazakhstan	Thu Jul 30, 2020 21:29 UTC	Proton-M/Briz-M	Progress MS & Progress ISS	StatusActive	415.0	Success	2020-07-30 21:29:00+00:00	2020
4	ULA	SLC-41, Cape Canaveral AFS, Florida, USA	Thu Jul 30, 2020 11:50 UTC	Atlas V 541	Perseverance	StatusActive	145.0	Success	2020-07-30 11:50:00+00:00	2020

```
[In]: df['year'].value_counts().reset_index()
```

	index	year
0	2021	120
1	2018	117
2	2017	114
3	2019	113
4	2020	113
...	...	...
56	2015	57
57	2022	52
58	1994	35
59	1958	28
60	1957	3

60 rows × 2 columns

```
[In]: fig = px.bar(df, x = "index", y = "year", title = "Missions Number by Year")
fig.show()
```



```
[In]: # Countries and Mission success
from sklearn.preprocessing import LabelEncoder
encoder = LabelEncoder()
df['Status Mission'] = encoder.fit_transform(df['Status Mission'])
encoder
```

```
[Out]:
LabelEncoder()
```

```
[In]: colors = {
                'USA': 'red',
                'China': 'blue',
                'Russian Federation': 'green',
                'Japan': 'brown',
                'Israel': 'red',
                'New Zealand': 'blue',
                'Iran': 'green',
                'France': 'red',
                'India': 'blue',
                'North Korea': 'green',
                'Pacific Ocean': 'red',
                'South Korea': 'blue',
                'Brazil': 'green',
                'Kenya': 'red',
                'Australia': 'blue'}
colors
```

```
[Out]:
{'USA': 'red',
 'China': 'blue',
 'Russian Federation': 'green',
 'Japan': 'brown',
 'Israel': 'red',
 'New Zealand': 'blue',
 'Iran': 'green',
 'France': 'red',
 'India': 'blue',
 'North Korea': 'green',
 'Pacific Ocean': 'red',
 'South Korea': 'blue',
 'Brazil': 'green',
 'Kenya': 'red',
 'Australia': 'blue'}
```

```
[In]: countries_dict = {
                'Russia': 'Russian Federation',
                'New Mexico': 'USA',
                'Pacific Ocean': 'USA',
                'General Mexico Sea': 'USA',
                'Gulfio Mexico Sea': 'USA',
                'Yellow Sea': 'China',
                'Shahrud Missile Test Site': 'Iran',
                'Pacific Ocean': 'USA'}
df['Country'] = df['Location'].str.split(', ').str[-1].replace(countries_dict)
df.head()
```

	Company Name	Location	Datum	Detail	Status Rocket	Rocket	Status Mission	date	year	country
0	Arianespace	Kourou	Fri Aug 07, 2020 05:12 UTC	Falcon 9 Block 5	Starlink V1 L9 & BlueOrig	StatusActive	811.0	Success	2020-08-07 05:12:00+00:00	2020	USA
1	CASC	Site 9401 (SLS-2), Jiuquan Satellite Launch Ce...	Thu Aug 06, 2020 04:01 UTC	Long March 2D	Gaofen-9 04 & Q-SAT	StatusActive	29.75	Success	2020-08-06 04:01:00+00:00	2020	China
2	SpaceX	Pad A, Boca Chica, Texas, USA	Thu Aug 06, 2020 07:37 UTC	Starship Prototype	150 Meter Hop	StatusActive	NaN	Failure	2020-08-06 07:37:00+00:00	2020	USA
3	Roscosmos	Site 200/39, Baikonur Cosmodrome, Kazakhstan	Thu Jul 30, 2020 21:29 UTC	Proton-M/Briz-M	Progress MS & Progress ISS	StatusActive	415.0	Success	2020-07-30 21:29:00+00:00	2020	Kazakhstan
4	ULA	SLC-41, Cape Canaveral AFS, Florida, USA	Thu Jul 30, 2020 11:50 UTC	Atlas V 541	Perseverance	StatusActive	145.0	Success	2020-07-30 11:50:00+00:00	2020	USA

```
[In]: from plotly.subplots import make_subplots
fig = make_subplots(rows = 4, cols = 4, subplot_titles = df['country'].unique())
for i, c in enumerate(df['country'].unique(), 1):
    counts = df[df['country'] == c].value_counts('Status Mission')
    fig.add_bar(x=counts.index, y=counts, name=c, marker_color=colors[c], row=(i-1)//4+1, col=(i-1)%4+1)
    fig.add_trace(go.Bar(x = counts.index, y = counts.values, name = c, marker = {'color': colors}), showlegend = False)
    fig.add_text(text=c, row = (i//4) + 1, col = (i%4)+1)
fig.update_layout(height=1000, width=1200, title_text="Countries and Mission Status", height = 1000, width = 1200)
fig.for_each(lambda t: t.update(textposition = "Outside"), row = 1, col = 1)
fig.show()
```