

```
data['Name'].value_counts()

University College of Engineering      7
Jamia Millia Islamia                  5
Siddaganga Institute of Technology    5
Delhi Technological University         5
National Institute of Technology Hamirpur 5
..
Government Engineering College        1
Government Engineering College, Thrissur 1
Amity University Gwalior              1
SR Engineering College                1
K.S.Rangasamy College of Technology    1
Name: Name, Length: 323, dtype: int64
```

```
data['Score'].value_counts()

33.58    3
36.42    3
33.98    3
33.24    3
35.89    3
..
40.24    1
39.97    1
39.93    1
39.68    1
36.21    1
Name: Score, Length: 698, dtype: int64
```

All the files for this competition are loaded ,as I am an engineering student I am interested in engineering dataset ,so I took engineering.csv file

```
data.shape

(800, 12)
```

This dataset has 800 observations and 12 features

```
data=pd.read_csv('../input/nirf-indian-universities-rankings-since-2017/Engineering.csv')
```

The 12 columns are 1)Institute ID 2)Name 3)City 4)State 5)Score 6)TLR 7)RPC 8)GO 9)OI 10)PERCEPTION 11)Year 12)Year_rank

```
data.shape
```

Here we can see the top5 rows of the dataset and all the 12rows of the dataset

There are no null values.

we can observe that some columns has numeric values and some are strings

```
data.columns
```

```
Index(['Institute ID', 'Name', 'City', 'State', 'Score', 'TLR', 'RPC',  
      'GO',  
      'OI', 'PERCEPTION', 'Year', 'Year_rank'],  
      dtype='object')
```

Here we can see the last 5 rows of the dataset.

we may confuse that top5 contains year 2021 and the last5 contains year 2017,i.e.,This NRIF rankings are from 2017 to 2021.

```
data.head(5)
```

	Institute ID	Name	City \
0	IR-E-U-0456	Indian Institute of Technology Madras	Chennai
1	IR-E-I-1074	Indian Institute of Technology Delhi	New Delhi
2	IR-E-U-0306	Indian Institute of Technology Bombay	Mumbai
3	IR-E-I-1075	Indian Institute of Technology Kanpur	Kanpur
4	IR-E-U-0573	Indian Institute of Technology Kharagpur	Kharagpur

	State	Score	TLR	RPC	GO	OI	PERCEPTION	Year
\								
0	Tamil Nadu	90.19	95.47	96.43	81.92	62.44	100.00	2021
1	Delhi	88.96	91.76	95.82	80.97	66.39	98.63	2021
2	Maharashtra	85.16	89.32	92.56	79.71	53.68	92.88	2021
3	Uttar Pradesh	83.22	86.71	83.13	89.79	55.96	87.05	2021
4	West Bengal	82.03	80.51	88.59	83.01	58.46	88.50	2021

	Year_rank
0	1
1	2
2	3
3	4
4	5

My dataset has 4 categorical columns and 8 numerical columns,In that 8 numerical columns 2 are integer and 6 are floating values

```
data.tail(5)
```

	Institute ID
Name \	
795	IR17-ENGG-2-12411
	Sri Sai Ram Engineering

College		
796	IR17-ENGG-2-12581	ST. Joseph's College of
Engineering		
797	IR17-ENGG-2-1-31	Andhra
University		
798	IR17-ENGG-2-1-2810997882	K.S.Rangasamy College of
Technology		
799	IR17-ENGG-2-1-365	National Institute of Technology
Meghalaya		

	City	State	Score	TLR	RPC	G0	OI
795	Chennai	Tamil Nadu	36.64	54.19	3.84	59.32	68.75
796	Chennai	Tamil Nadu	36.51	59.47	6.74	49.39	64.92
797	Visakhapatnam	Andhra Pradesh	36.42	41.94	15.81	66.86	48.27
798	Tiruchengode	Tamil Nadu	36.40	53.78	14.29	46.44	60.98
799	Shillong	Meghalaya	36.21	62.02	4.12	48.34	60.60

	PERCEPTION	Year	Year_rank
795	4.86	2017	96
796	2.80	2017	97
797	9.03	2017	98
798	5.85	2017	99
799	6.29	2017	100

```
data.dtypes.value_counts()
```

```
float64    6
object     4
int64      2
dtype: int64
```

There are no null values in the dataset

eventhough we can't find null values in top5 and last5rows,just for classification I verified nullvalues.

```
missing=data.isnull().sum()
```

Here the mean is 87.9 and the median is 80.5.

There is no much difference between mean and median,i.e.,no outliers

```
missing
```

```
Institute ID    0
Name            0
City            0
State           0
Score           0
TLR             0
RPC             0
GO              0
OI              0
PERCEPTION      0
Year            0
Year_rank       0
dtype: int64
```

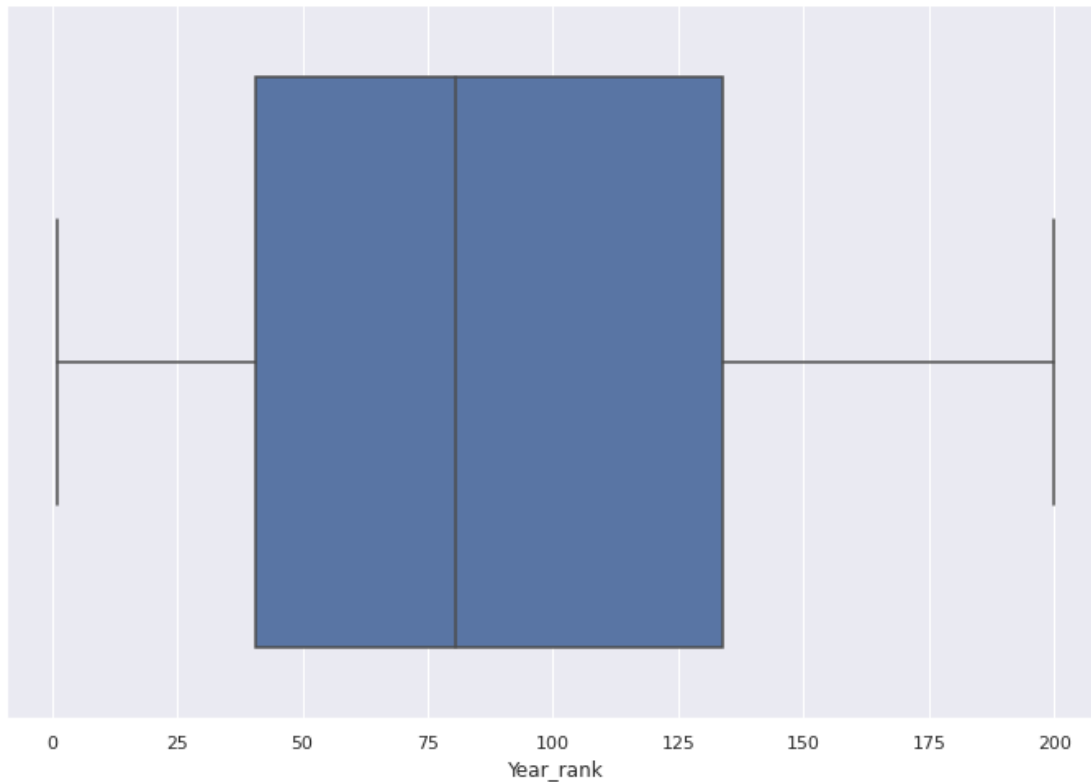
This returns different statistics like count of values , unique values, top and frequency of occurrences. The count values are 800, unique values are 323 , frequency of occurrences are 7

```
sns.boxplot(data['Year_rank'])
```

```
/opt/conda/lib/python3.7/site-packages/seaborn/_decorators.py:43:
FutureWarning: Pass the following variable as a keyword arg: x. From
version 0.12, the only valid positional argument will be `data`, and
passing other arguments without an explicit keyword will result in an
error or misinterpretation.
```

```
FutureWarning
```

```
<AxesSubplot:xlabel='Year_rank'>
```



HISTOGRAM

```
data['Year_rank'].describe()
```

```
count      800.000000
mean       87.941250
std        56.380327
min         1.000000
25%        40.750000
50%        80.500000
75%       134.000000
max       200.000000
Name: Year_rank, dtype: float64
```

If the score is more, ranking will be less. so the top universities score is in between 60 and 90

```
data['Name'].describe()
```

```
count      800
unique     323
top      University College of Engineering
freq              7
Name: Name, dtype: object
```

Here in this plot, peak value is in between 30 to 40 and also the tail part says that very few universities have 70 to 100

```
data['Name'].unique()
```

```
array(['Indian Institute of Technology Madras',  
      'Indian Institute of Technology Delhi',  
      'Indian Institute of Technology Bombay',  
      'Indian Institute of Technology Kanpur',  
      'Indian Institute of Technology Kharagpur',  
      'Indian Institute of Technology Roorkee',  
      'Indian Institute of Technology Guwahati',  
      'Indian Institute of Technology Hyderabad',  
      'National Institute of Technology Tiruchirappalli',  
      'National Institute of Technology Karnataka',  
      'Indian Institute of Technology (Indian School of Mines)',  
      'Vellore Institute of Technology',  
      'Indian Institute of Technology Indore',  
      'Indian Institute of Technology (Banaras Hindu University)  
Varanasi',  
      'Institute of Chemical Technology', 'Amrita Vishwa  
Vidyapeetham',  
      'Jadavpur University', 'Anna University',  
      'Indian Institute of Technology Ropar',  
      'National Institute of Technology Rourkela',  
      'Indian Institute of Technology Patna',  
      'Indian Institute of Technology Gandhinagar',  
      'National Institute of Technology Warangal',  
      'Thapar Institute of Engineering and Technology',  
      'National Institute of Technology Calicut',  
      'Birla Institute of Technology & Science -Pilani',  
      'Indian Institute of Engineering Science and Technology',  
      'Indian Institute of Technology Bhubaneswar',  
      'National Institute of Technology Durgapur',  
      'Visvesvaraya National Institute of Technology',  
      'Amity University', 'Siksha `O` Anusandhan',  
      'Jamia Millia Islamia',  
      'S.R.M. Institute of Science and Technology',  
      'Aligarh Muslim University', 'Delhi Technological University',  
      'Malaviya National Institute of Technology',  
      'Shanmugha Arts Science Technology & Research Academy',  
      'Kalinga Institute of Industrial Technology',  
      'Indian Institute of Space Science and Technology',  
      'Indian Institute of Technology Mandi',  
      'Motilal Nehru National Institute of Technology',  
      'Indian Institute of Technology Jodhpur',  
      'National Institute of Technology Kurukshetra',  
      'Sri Sivasubramaniya Nadar College of Engineering',  
      'Birla Institute of Technology',  
      'Sardar Vallabhbhai National Institute of Technology',  
      'National Institute of Technology Silchar',  
      'Dr. B. R. Ambedkar National Institute of Technology',  
      'Koneru Lakshmaiah Education Foundation University (K L College
```

of Engineering)',
 'Manipal Institute of Technology', 'College of Engineering,
 Pune',
 'PSG College of Technology',
 'International Institute of Information Technology Hyderabad',
 'Sathyabama Institute of Science and Technology',
 'Kalasalingam Academy of Research and Higher Education',
 'Visvesvaraya Technological University',
 'Defence Institute of Advanced Technology',
 'National Institute of Technology Meghalaya',
 'Maulana Azad National Institute of Technology',
 'Chandigarh University',
 'Jawaharlal Nehru Technological University',
 'Indraprastha Institute of Information Technology, Delhi',
 'National Institute of Technology Raipur',
 'M. S. Ramaiah Institute of Technology',
 'Lovely Professional University', 'Banasthali Vidyapith',
 'Pandit Deendayal Petroleum University',
 'National Institute of Technology Srinagar',
 'Punjab Engineering College (Deemed to be University),
 Chandigarh',
 'Thiagarajar College of Engineering',
 'National Institute of Technology Patna',
 'Indian Institute of Information Technology Guwahati',
 'College of Engineering (A)', 'Graphic Era University',
 'International Institute of Information Technology Bangalore',
 'R.V. College of Engineering',
 'Sri Krishna College of Engineering and Technology',
 'Shri Mata Vaishno Devi University',
 'Pandit Dwarka Prasad Mishra Indian Institute of Information
 Technology, Design and Manufacturing (IIITDM) Jabalpur',
 'Indian Institute of Food Processing Technology (IIFPT)',
 'Veermata Jijabai Technological Institute, (VJTI, Mumbai)',
 'PES University', 'Manipal University Jaipur',
 'National Institute of Technology Goa',
 'National Institute of Technology, Jamshedpur',
 'Indian Institute of Information Technology Allahabad',
 'Netaji Subhas University of Technology (NSUT)',
 'Panjab University', 'Kumaraguru College of Technology',
 'University of Petroleum and Energy Studies',
 'National Institute of Technology Agartala',
 'Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science
 and Technology',
 'Jaypee Institute of Information Technology',
 'College of Engineering Trivandrum',
 'Bharati Vidyapeeth Deemed University College of Engineering',
 'The Northcap University', 'B.M.S. College of Engineering',
 'National Institute of Technology Hamirpur',
 'C.V. Raman Global University, Odisha',
 'Siddaganga Institute of Technology',

'Coimbatore Institute of Technology',
 'Shoolini University of Biotechnology and Management Sciences',
 'Tezpur University',
 'Karunya Institute of Technology and Sciences',
 'Atal Bihari Vajpayee Indian Institute of Information
 Technology and Management',
 'Dhirubhai Ambani Institute of Information and Communication
 Technology',
 'Guru Gobind Singh Indraprastha University',
 'Sri Sairam Engineering College',
 'Rajalakshmi Engineering College',
 'Vignan's Foundation for Science, Technology and Research",
 'B. S. Abdur Rahman Crescent Institute of Science and
 Technology',
 'Chitkara University', 'National Institute of Technology
 Manipur',
 'National Institute of Foundry and Forge Technology (NIFFT)',
 'Veer Surendra Sai University of Technology',
 'Maharaja Sayajirao University of Baroda',
 'Manav Rachna International Institute of Research & Studies',
 'Shri Ramdeobaba College of Engineering and Management',
 'Punjab Technical University',
 'Saveetha Institute of Medical and Technical Sciences',
 'Vishwakarma Institute of Technology',
 'University College of Engineering',
 'Mepco Schlenk Engineering College',
 'Vels Institute of Science, Technology & Advanced Studies
 (VISTAS)',
 'Dayalbagh Educational Institute',
 'YMCA University of Science & Tech (Formerly YMCA Institute of
 Engineering)',
 'Government College of Technology',
 'Jaypee University of Information Technology',
 'G. H. Rasoni College of Engineering',
 'PSG Institute of Technology and Applied Research',
 'University College of Engineering',
 'Chaitanya Bharathi Institute of Technology', 'SR University',
 'Vallurupalli Nageswara Rao Vignana Jyothi Institute of
 Engineering and Technology',
 'The Rashtrasant Tukadoji Maharaj Nagpur University',
 'Sri Krishna College of Technology',
 'NMAM Institute of Technology', 'P E S College of Engineering',
 'Army Institute of Technology',
 'New Horizon College of Engineering',
 'Vardhaman College of Engineering',
 'Bharatiya Vidya Bhavan's Sardar Patel Institute of
 Technology',
 'Pondicherry Engineering College',
 'National Institute of Technology Puducherry', 'Anurag
 University',

'Chandigarh Engineering College',
 'National Institute of Food Technology, Entrepreneurship &
 Management',
 'Yeshwantrao Chavan College of Engineering',
 'Dr. Vishwanath Karad MIT World Peace University',
 'SVKM's Narsee Monjee Institute of Management Studies',
 'BMS Institute of Technology & Management', 'Nirma University',
 'Amity University, Gwalior', 'CVR College Of Engineering',
 'Government Engineering College, Thrissur',
 'Jain University, Bangalore',
 'Maulana Abul Kalam Azad University of Technology',
 'Institute of Aeronautical Engineering',
 'National Institute of Technology Arunachal Pradesh',
 'Indira Gandhi Delhi Technical University for Women',
 'Sant Longowal Institute of Engineering & Technology',
 'Silicon Institute of Technology (SIT), Bhubaneswar',
 'Kongu Engineering College',
 'Goka Raju Ranga Raju Institute of Engineering & Technology',
 'Annamalai University',
 'Rajiv Gandhi Institute of Petroleum Technology',
 'Nitte Meenakshi Institute of Technology',
 'Institute of Engineering & Management',
 'R.M.K. Engineering College',
 'Noida Institute of Engineering & Technology',
 'Hindustan Institute of Technology and Science (HITS)',
 'Dr. D. Y. Patil Institute of Technology',
 'Guru Nanak Dev University', 'Sri Ramakrishna Engineering
 College',
 'KIET Group of Institutions', 'Galgotias University',
 'Velagapudi Ramakrishna Siddhartha Engineering College',
 'Sharda University', 'KLE Technological University',
 'Shri G. S. Institute of Technology & Science',
 'North Eastern Regional Institute of Science & Technology',
 'St. Josephs College of Engineering',
 'Lakshmi Narain College of Technology',
 'Centurion University of Technology and Management',
 'Rabindranath Tagore University',
 'National Institute of Technology Uttarakhand',
 'JSS Science and Technology University',
 'Easwari Engineering College',
 'International Institute of Information Technology
 Bhubaneswar',
 'National Engineering College', 'Sona College of Technology',
 'C M R Institute of Technology',
 'G.L.Bajaj Institute of Technology and Management',
 'DIT University', 'The National Institute of Engineering',
 'K. J. Somaiya College of Engineering',
 'Kakatiya Institute of Technology & Science',
 'Walchand College of Engineering', 'Sri Venkateswara
 University',

'Indian Institute of Technology (BHU) Varanasi',
 'Amrita School of Engineering',
 'Thapar Institute of Engineering & Technology',
 'Birla Institute of Technology & Science',
 'Amity University Noida',
 'S. R. M. Institute of Science and Technology',
 'College of Engineering Pune',
 'Indraprastha Institute of Information Technology Delhi',
 'Koneru Lakshmaiah Education Foundation University',
 'Punjab Engineering College (Deemed To Be University)',
 'College of Engineering(A)', 'R. V. College of Engineering',
 'Veermata Jijabai Technological Institute',
 'National Institute of Technology Jamshedpur',
 'Indian Institute of Information Technology Design &
 Manufacturing Jabalpur',
 'C.V. Raman Global University', 'Jain University',
 'Vignan's Foundation for Science, Technology & Research',
 'YMCA University of Science and Technology',
 'National Institute of Food Technology, Enterprenurship &
 Management',
 'Dayananda Sagar College of Engineering',
 'N M A M Institute of Technology',
 'Heritage Institute of Technology', 'R. M. K. Engineering
 College',
 'Alliance University', 'SR Engineering College',
 'Amity University Gwalior', 'Haldia Institute of Technology',
 'Government Engineering College',
 'Harcourt Butler Technical University',
 'Maharshi Karve Stree Shikshan Samstha's Cummins College of
 Engineering for Women',
 'ManavRachna International Institute of Research & Studies',
 'Sri Venkateswara College of Engineering',
 'School of Engineering, Cochin University of Science and
 Technology',
 'Anurag Group of Institutions',
 'Indian Institute of Information Technology, Design &
 Manufacturing',
 'Madan Mohan Malaviya University of Technology',
 'Sree Vidyanikethan Engineering College',
 'JNTUA College of Engineering', 'BNM Institute of Technology',
 'Vasavi College of Engineering',
 'Gayatri Vidya Parishad College of Engineering',
 'G.Pulla Reddy Engineering College',
 'Sri Sai Ram Institute of Technology',
 'Dharmsinh Desai University',
 'The LNM Institute of Information Technology',
 'Guru Jambheshwar University of Science and Technology',
 'Pimpri Chinchwad College of Engineering',
 'Ramrao Adik Institute of Technology', 'BVRIT Hyderabad',
 'Indian Institute of Technology (Indian School of Mines)

Dhanbad',
 'Thapar Institute of Engineering and Technology (Deemed-to-be-university)',
 'SRM Institute of Science and Technology',
 'College of Engineering',
 'Malviya National Institute of Technology',
 'Indraprastha Institute of Information Technology',
 'College of Engineering Visakhapatnam',
 'National Institute of Industrial Engineering',
 'B. M. S. College of Engineering', 'PEC University of Technology',
 'C. V. Raman College of Engineering',
 'Hindustan Institute of Technology and Science',
 'Bannari Amman Institute of Technology',
 'Vignan's Foundation for Science, Technology and Research',
 'Tamil Nadu Veterinary & Animal Sciences University',
 'Jain university', 'Sri Jayachamarajendra College of Engineering',
 'Saveetha Engineering College',
 'N. M. A. M. Institute of Technology',
 'CVR College of Engineering',
 'G. B. Pant University of Agriculture and Technology',
 'NITTE Meenakshi Institute of Technology',
 'YMCA University of Science & Tech',
 'PSNA College of Engineering and Technology',
 'Bansilal Ramnath Agarwal Charitable Trust's Vishwakarama Institute of Technology',
 'Narula Institute of Technology',
 'Dr. Ambedkar Institute of Technology',
 'Padmasri Dr. B.V. Raju Institute of Technology',
 'Sri Venkateswara College of Engineering and Technology',
 'G. Pulla Reddy Engineering College',
 'Sardar Patel College of Engineering',
 'Shri Vishnu Engineering College for Women',
 'Dr. Mahalingam College of Engineering and Technology',
 'Maharshi Karve Stree Shikshan Samstha's Cummins College of Engineering for Women',
 'Sagi Ramakrishnam Raju Engineering College',
 'Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College',
 'Basaveshwar Engineering College',
 'Thakur College of Engineering & Technology',
 'Shri Vile Parle Kelavani Mandal's Dwarkadas J. Sanghvi College of Engineering',
 'M. G. R. Educational and Research Institute',
 'Dr. B. R. Ambedkar Institute of Technology',
 'National Institute of Technology Surathkal',
 'Indian Institute of Engineering Science and Technology, Shibpur',
 'International Institute of Information Technology',

```

        'AU College of Engineering',
        'B. S. Abdur Rahman Institute of Science and Technology',
        'Indian Institute of Engineering Science And Technology,
Shibpur',
        'Shanmugha Arts Science Technology & Research Academy
(SASTRA)',
        'Thapar University',
        'Sri Sivasubramaniya Nadar College of Engineering',
        'Indian Institute of Technology (Banaras Hindu University),
Varanasi',
        'S.R.M Institute of Science and Technology',
        'National Institute of Industrial Engineering, Mumbai',
        'Zakir Husain College of Engineering & Technology',
        'Visvesvaraya National Institute of Technology, Nagpur (Deemed
University)',
        'AMITY University',
        'Indraprastha Institute of Information Technology-New Delhi
(IIIT New Delhi, Delhi)',
        'Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and
Technology',
        'Karunya Institute of Technology and Sciences-Coimbatore',
        'G. H. Rasoni College of Engineering, NAGPUR',
        'Pandit Dwarka Prasad Mishra Indian Institute of Information
Technology, Design and Manufacturing (IIITDM), Jabalpur',
        'Malaviya National Institute of Technology, Jaipur',
        'Faculty of Technology & Engineering, The Maharaja Sayajirao
University of Baroda',
        'PSNA College of Engineering and Technology, Dindigul',
        'Jaypee University of Information Technology-Solan',
        'Punjab Engineering College',
        'B.S. Abdur Rahman Institute of Science and Technology',
        'C.V. Raman College of Engineering',
        'Maharashtra Institute of Technology, Pune',
        'Sri Sai Ram Engineering College',
        "ST. Joseph's College of Engineering", 'Andhra University',
        'K.S.Rangasamy College of Technology'], dtype=object)

```

```
data['Name'].nunique()
```

323

SCATTER PLOT

scatter plots use a collection of points placed using Cartesian coordinates to display values from two variables. By displaying a variable in each axis, we can detect if a relationship or correlation between the two variables exists. Scatterplots are also great for observing the spread of the data as they retain the exact data values and sample size.

```
sns.scatterplot(x='TLR', y='Year_rank', data=data)
```

```
<AxesSubplot:xlabel='TLR', ylabel='Year_rank'>
```

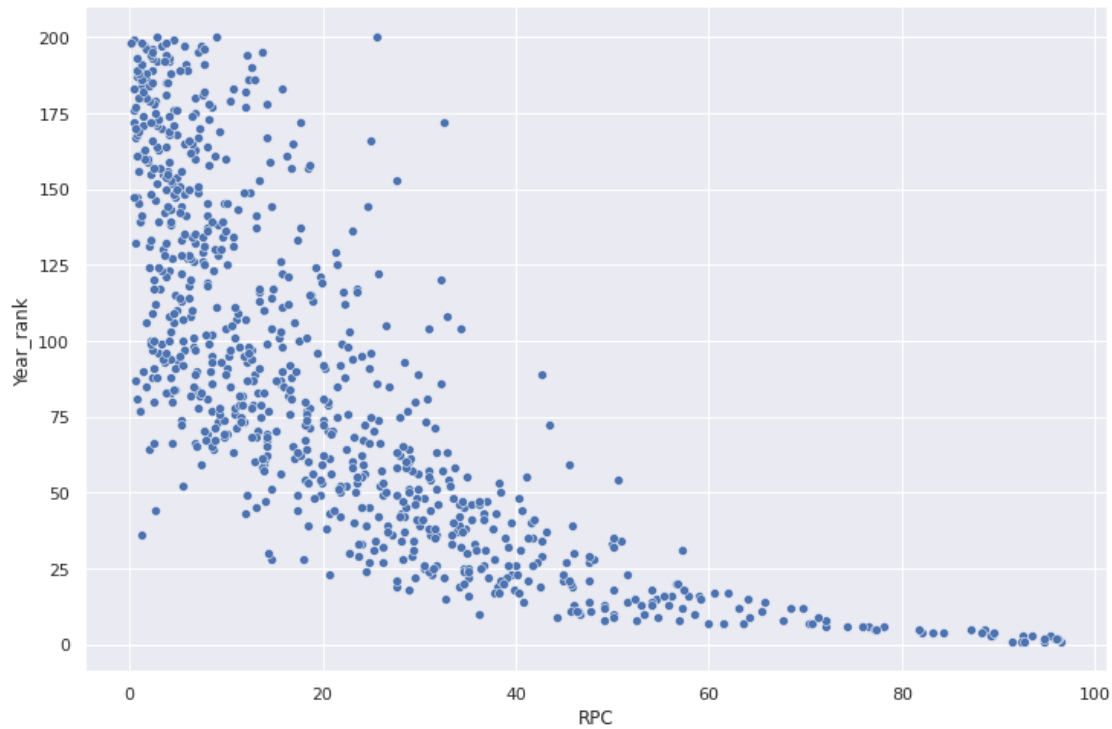


scatterplot is drawn TLR on xaxis and Year_rank on yaxis

Here we can clearly observe that if the TLR(Teaching, Learning and Resources) is high ,rank is low. and only few universities have best TLR in range 70 to 100

```
sns.scatterplot(x='RPC', y='Year_rank', data=data)
```

```
<AxesSubplot:xlabel='RPC', ylabel='Year_rank'>
```

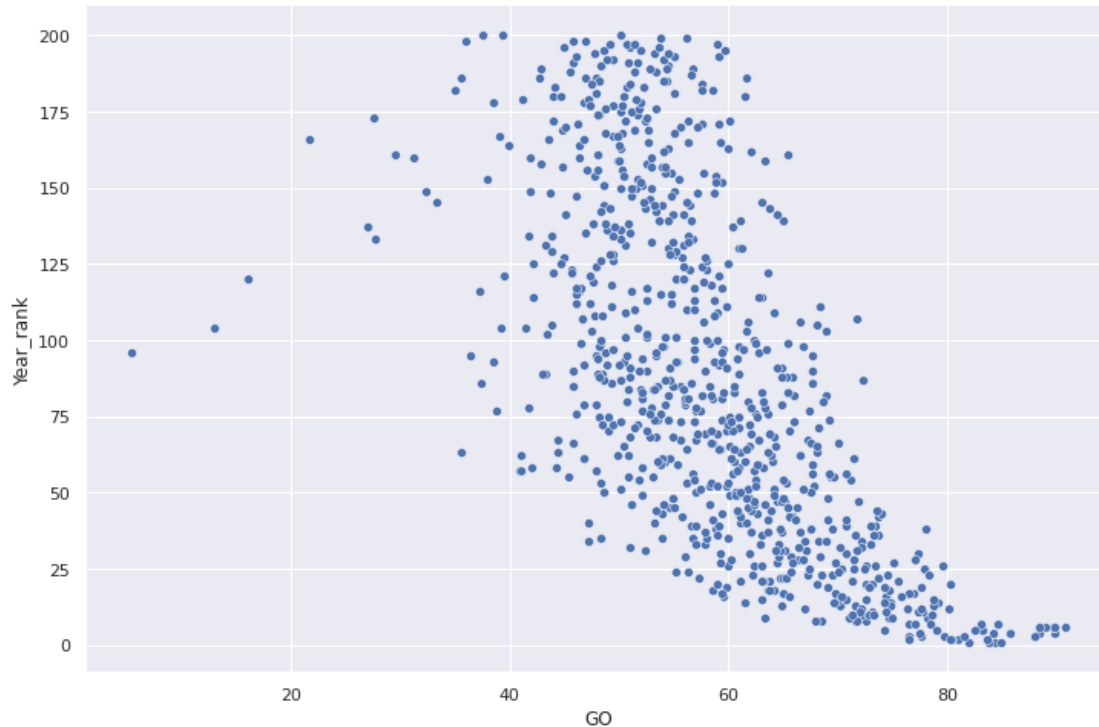


scatterplot is drawn between RPC on Xaxis and Year_rank on Yaxis

Here we can clearly observe that if the RPC(Research&ProffessionalPractice) is more ,rank is less. and only few universities have best RPC in the range of 60 to 100.

```
sns.scatterplot(x='G0', y='Year_rank', data=data)
```

```
<AxesSubplot:xlabel='G0', ylabel='Year_rank'>
```

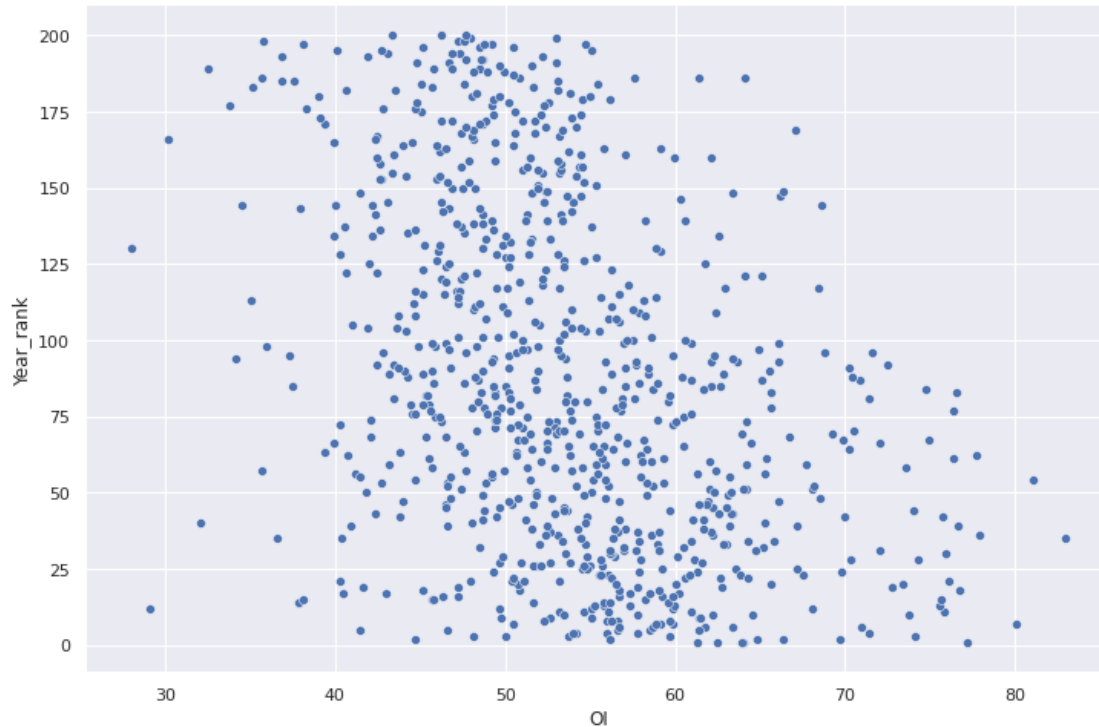


scatterplot is drawn between GO on xaxis and Year_rank on yaxis

Here we can clearly observe that if the GO(Graduation Outcomes) is more ,rank is less. and only few universities have more GO in range 70 to 100

```
sns.scatterplot(x='GO', y='Year_rank', data=data)
```

```
<AxesSubplot:xlabel='GO', ylabel='Year_rank'>
```

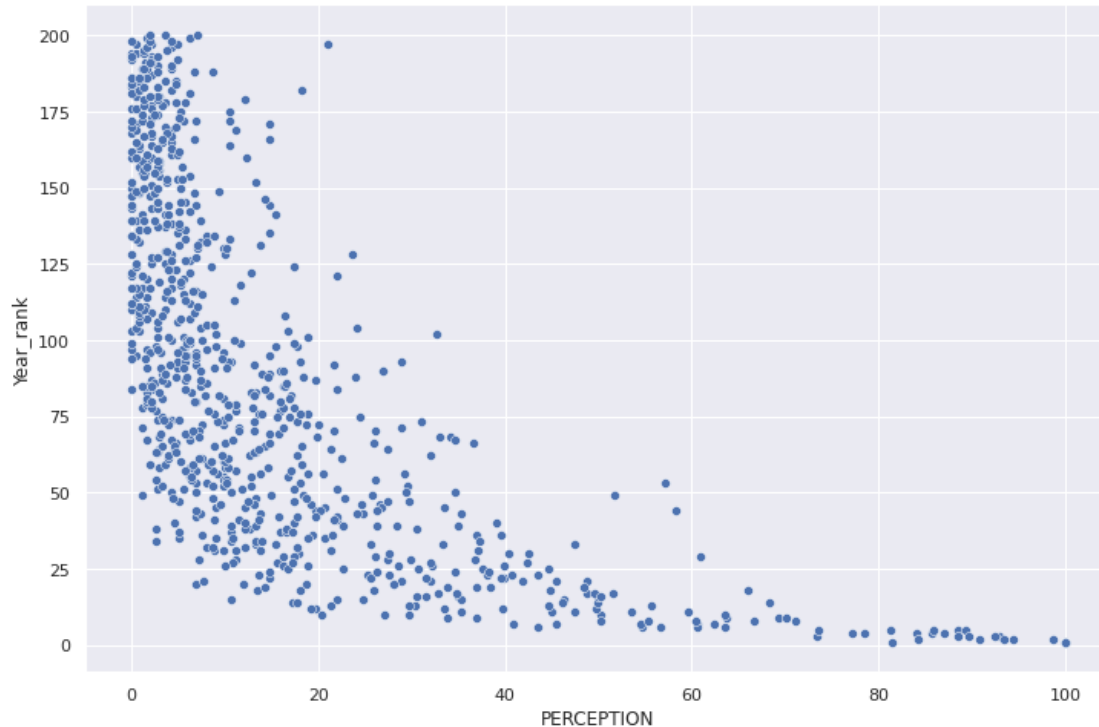


scatterplot is drawn between OI on xaxis and Year_rank on yaxis

Here we can clearly observe that if the OI(outreach&Inclusivity) is more, rank is less.and only few universities in the range of 60 to 100

```
sns.scatterplot(x='PERCEPTION', y='Year_rank', data=data)
```

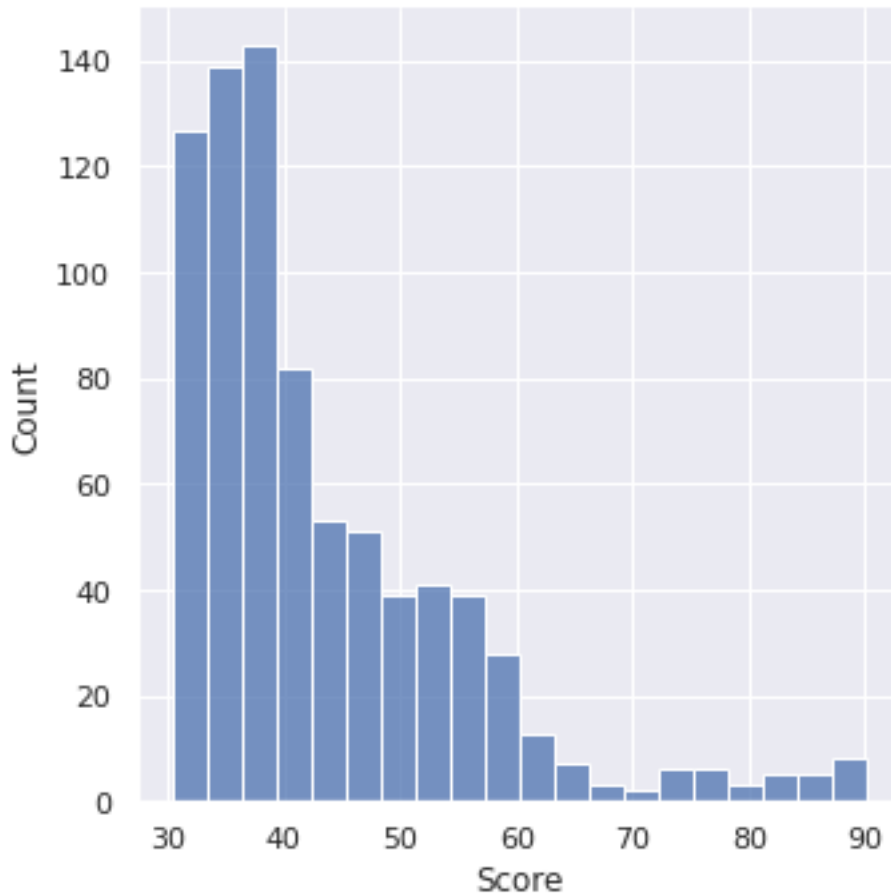
```
<AxesSubplot:xlabel='PERCEPTION', ylabel='Year_rank'>
```

scatterplot is drawn between PERCEPTION on Xaxis and Year_rank on Yaxis

Here we can clearly observe that if the PERCEPTION is high ,rank is less.and only few universities in range of 40 to 100

```
sns.set(rc={'figure.figsize':(12,8)})  
sns.displot(data['Score'], kde=False, bins=20) ;
```



Scatterplot is drawn between score on xaxis and Year_rank on yaxis.

Here we can clearly observe that if the score is more,rank is less.and only few universities in range of 60 to 90

Here numeric columns are stored in variable called number if features and columns are extracted.We can see that nearly 8 columns outof 12 have numerical values.

They are1)Score 2)TLR 3)RPC 4)GO 5)OI 6)PERCEPTION 7)Year 8)Year_rank

CORRELATION

correlation can be calculated only on numerical columns , we can't calculate correlation on non-numeric.

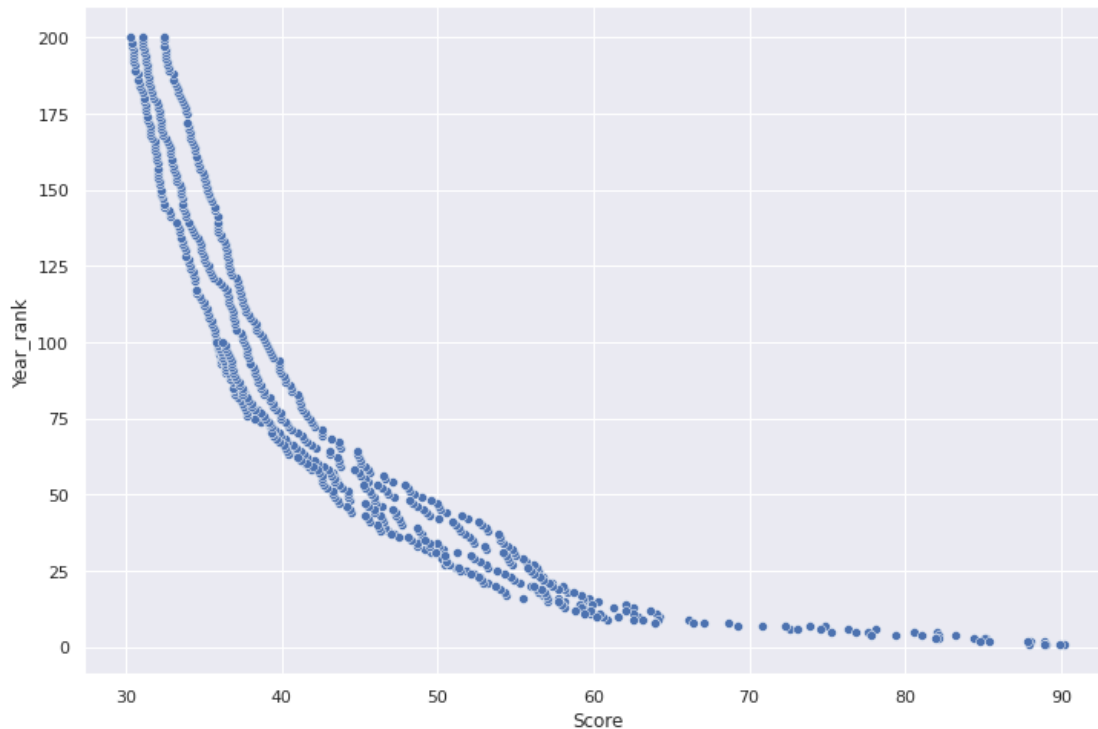
Here numeric columns are extracted excluding strings.

Here numeric columns are extracted excluding strings.

It gave rows and columns .And by this we can conclude that 8columns are numeric outof 12.

```
sns.scatterplot(x='Score', y='Year_rank', data=data)
```

<AxesSubplot:xlabel='Score', ylabel='Year_rank'>



It gives all the string columns. We have nearly 4 out of 12 columns

They are 1) InstituteID 2) Name 3) City 4) State

```
numeric_features = data.select_dtypes(include = [np.number])
numeric_features.columns
```

```
Index(['Score', 'TLR', 'RPC', 'G0', 'OI', 'PERCEPTION', 'Year',
       'Year_rank'], dtype='object')
```

It gave rows and columns. By this we can conclude that 4 string values are present

```
numeric_features1 = data.select_dtypes(exclude = ['0'])
numeric_features1.columns
```

```
Index(['Score', 'TLR', 'RPC', 'G0', 'OI', 'PERCEPTION', 'Year',
       'Year_rank'], dtype='object')
```

To find the correlation, we are using corr method.

```
numeric_features.shape, numeric_features1.shape
((800, 8), (800, 8))
```

This is the correlation matrix for all the 9 numerical columns

```
categorical_features = data.select_dtypes(include = [np.object])
categorical_features.columns
```

```
Index(['Institute ID', 'Name', 'City', 'State'], dtype='object')
```

We can conclude that top5 numerical columns are Year_rank , Year , OI , PERCEPTION , G0

```
categorical_features.shape
```

```
(800, 4)
```

It is the boxplot on Year_rank and we can observe that median is in between 75 to 80.there are outliers.

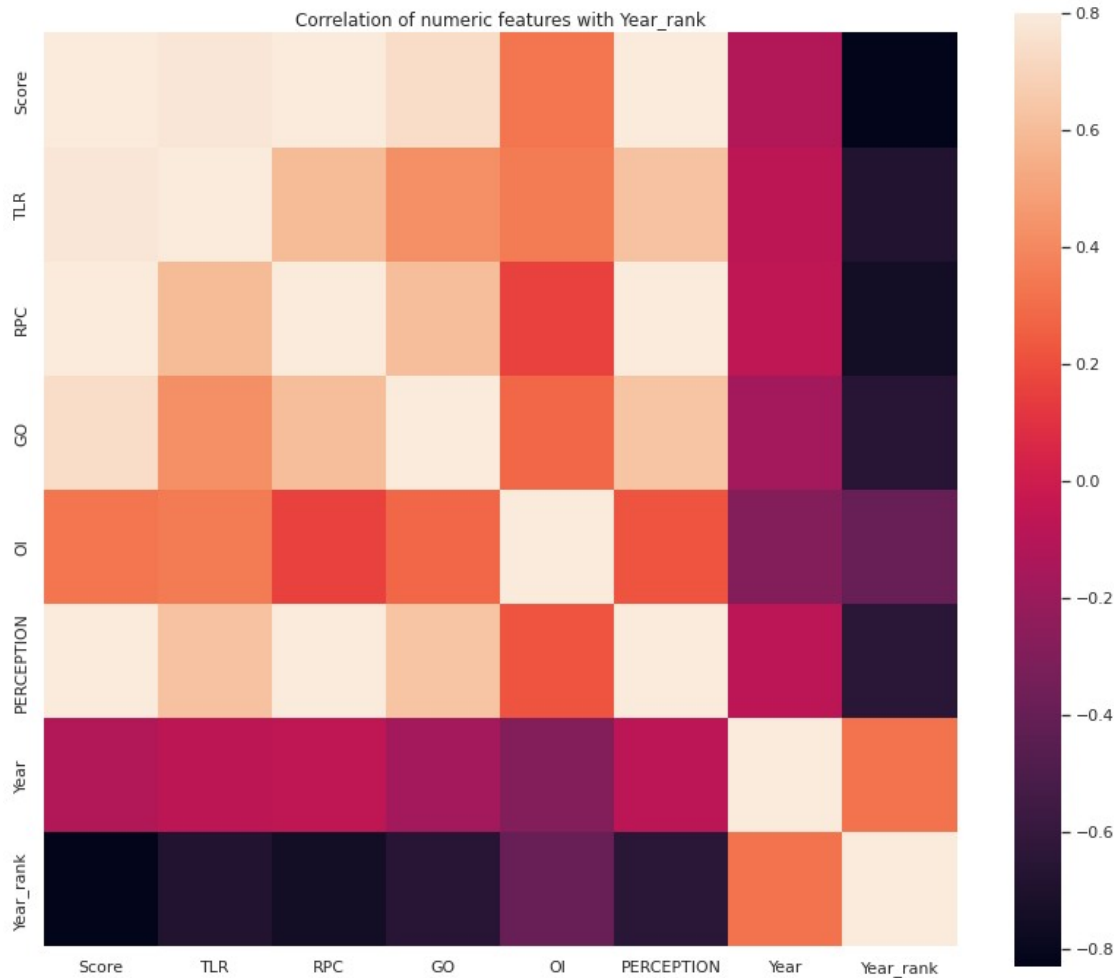
```
correlation = numeric_features.corr()
print(correlation['Year_rank'].sort_values(ascending = False), '\n')
```

```
Year_rank      1.000000
Year           0.316048
OI            -0.397382
PERCEPTION    -0.650370
G0            -0.657566
TLR           -0.678113
RPC           -0.754365
Score         -0.830824
Name: Year_rank, dtype: float64
```

We can observe the array of all the Universities names that has unique names

```
f, ax = plt.subplots(figsize = (14, 12))
plt.title('Correlation of numeric features with Year_rank')
sns.heatmap(correlation, square=True, vmax=0.8)
```

```
<AxesSubplot:title={'center': 'Correlation of numeric features with
Year_rank'}>
```



We can observe nearly 323 universities have unique names

```
k=5
cols = correlation.nlargest(k, 'Year_rank')['Year_rank'].index
print(cols)

Index(['Year_rank', 'Year', 'OI', 'PERCEPTION', 'GO'], dtype='object')
```

It gives value counts for all the university names. we can observe that University college of Engineering has the highest count that is 7 and the least count is 1

conclusion

Finally I came up with all the solutions to the questions in your mind by performing Exploratory Data Analysis on Nri rankings of universities.

- Educational Institute or University with high Nri ranking can get more student admission interest.
- Nri ranking will help student to choose the best University.

- Nrif ranking is based on five terms: 1.Teaching, Learning and Resources(TLR) 2.Research & Professional Practice(RP) 3.Graduation Outcomes(GO) 4.Outreach & Inclusivity(OI) 5.Perception
- If TLR , RP , GO and OI are good and high ,in the range of 70 to 100 then their Nrif rank will be less
- One should get the good scores range of 70 to 100 to get into best Nrif ranking Universities.
- In this dataset there are 800 rows and 12 columns , and I have obtained clearcut graphs on the five parameters vs ranking.
- There are 323 Universities in Number.

Self-belief and Hardwork will always earn you Success.

- I hope the notebook was insightful , and as I am new to kaggle ,I would really appreciate some feedback:-)