

School of Computing Science and Engineering

Lab exercise

Code/Course	:	CSE3020 – Data Visualization	Date	:	23/11/2021
Lab Experiment		Text data visualization using Python	Slot	:	L9+L10

Pre-requisite: We will assume you are moderately familiar with basic concepts in Python

Dataset : Airline, Airport and Route datasets

Reference : (<https://www.dataquest.io/blog/python-data-visualization-libraries/>)

Exercises

- **Pre Processing : Assign Column Headers to the given datasets**

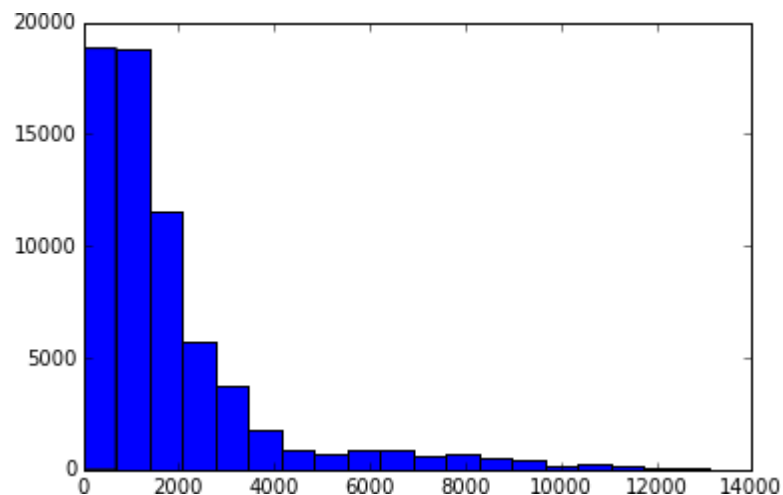
	i d	name	city	count ry	cod e	icao	latitu de	longitu de	altitu de	offs et	d st	timezone
01		Goro ka	Goro ka	Papu a New Guin ea	GK A	AYG A	- 6.081 689	145.391 881	5282	10	U	Pacific/Port_ M oresby
12		Mada ng	Mada ng	Papu a New Guin ea	MA G	AY MD	- 5.207 083	145.788 700	20	10	U	Pacific/Port_ M oresby

	id	name	alias	iata	icao	callsign	country	active	
0	1	Private flight	\N	-	NaN	NaN	NaN	NaN	

	airline	airline_id	source	source_id	dest	dest_id	codeshare	stops	equipment
0	2B	410	AER	2965	KZN	2990	NaN	0	CR2
1	2B	410	ASF	2966	KZN	2990	NaN	0	CR2

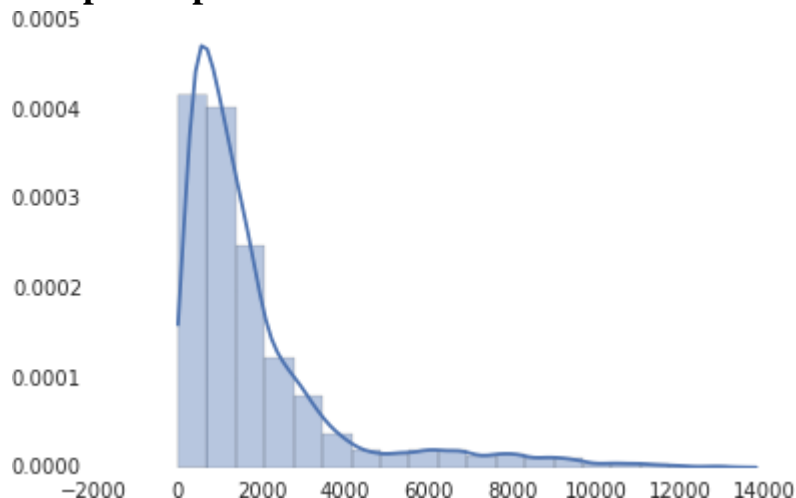
- Make histogram for route length, bin the values into ranges and count how many routes fall into each range

Sample output



- Use seaborn for route dataset (route Length)

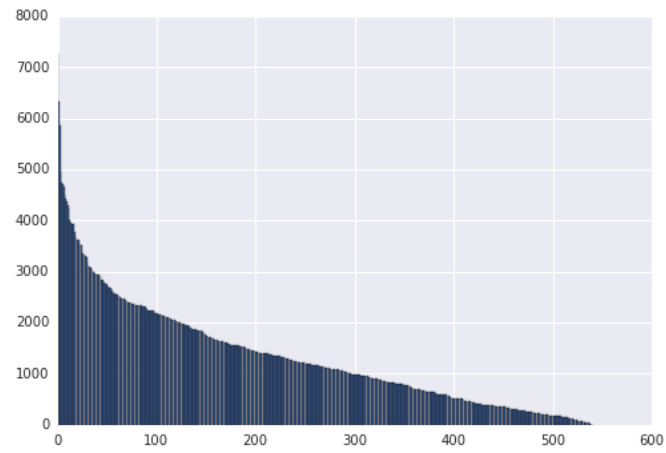
Sample output



- Bar chart - plot each airline against the average route length each airline

flies

Sample output



- **Create a scatter plot comparing the airline ids to the name lengths**

