DATA ANALYSIS ON UDEMY COURSES – DATA PROJECT 8

1. Import required libraries

2. Read the data file

3. Check data types of each column

4. You will find published_timestamp column of object data type, we need to convert this into date time datatype (it should be converted for the whole dataset , try to convert while reading the csv file)

5. Check data types again

6. Display top 10 rows of the dataset

7. Check last 5 rows of the dataset

8. Find shape of our dataset (number of rows and number of columns)

9. Getting information about our dataset like total number rows, total number of columns, datatypes of each column and memory requirement

10. Check null values in the dataset

    a. Try to find out if there are any missing values in the dataset

    b. Create heatmap as well

11. Check for duplicate data and drop them

12. Find out number of courses per subjects

    a. Elaborate your observation as well in comments

    b. You need to visualize data using barplot which is having xlabel as 'subjects' and ylabel as 'number of courses per subject'

    c. Take care of not overlapping words

13. For which levels, udemy courses providing the courses

a. Kindly visualize the data using barplot

14. Display the count of paid and free courses

a. Visualize it as well

15. Which course has more lectures (free or paid)?

a. A barplot with this looks great

16. Which courses have a higher number of subscribers free or paid?

a. Barplot will work here too

17. Which level has the highest number of subscribers?

18. Find most popular course title

19. Display 10 most popular courses as per number of subscribers

20. Find the course which is having the highest number of reviews.

21. Does price affect the number of reviews?

a. Use scatterplot for visualizing it

b. No overlapping should be there

22. Find total number of courses related to python

a. Elaborate your observation

23. Display 10 most popular python courses as per number of subscribers

a. Would be great practice is you use barplot to visualize it

24. In which year the highest number of courses were posted?

a. Cause you have to show year, try to use countplot to visualize it

25. Display category-wise count of posted subjects [year wise]

Bonus – Try to add some more visualization , can add as many other points as well.