

Question Bank for DSBDL

- 1)What is Pandas library in Python?
- 2)List some key features of Pandas.
- 3)What is Numpy Library in Python?
- 4)What is matplotlib library?
- 5)What is difference between seaborn library and matplotlib?
- 6)Is Sk learn and Sci kit Learn are same library?what is use of it in data science.
- 7)What are functions in Pandas and NumPy library?
- 6)What is data frame in Python?
- 7)How to find duplicates in python?(Python Command)
- 8)what is use of describe command?
- 10)Which are Naive Bayes classification algorithm used in python?
- 11)What is significance of Confusion matrix?
- 12)What is TP,TN,FP,FN in confusion matrix?
- 13)What is recall?
- 14)What is precision?
- 15)What is f1 score?
- 16)What is need of data visualization in data science?
- 17)What is outlier?
- 18)When to use histogram and pie chart?
- 19)What are the challenges in Big data visualization?
- 20)What is joint plot,dist plot?
- 21)What are tools used for data visualization?
- 22)What is data Wrangling?
- 23)What is data transformation?
- 24)What is use of standard scalar function in python?
- 25)What is Hadoop?
- 26)What is HDFS and Map reduce?
- 27)What are the components of Hadoop Ecosystem?
- 26)What is Scala?
- 27)What are features of Scala?
- 28)How scala is different from java?
- 29)List applications of Scala.
- 30)What is data science?
- 31)What is Big data?
- 32)What are the characteristics of Big data?
- 33)List phases in data science life cycle?
- 33)What is Central tendency?
- 34)What is dispersion?
- 35)What is mean,mode,median?Calculate same for
10,22,13,10,21,43,77,21,10
- 36)What is Variance?
- 37)What is Standard Deviation?
- 38)What is meant by posterior probability in naive Bayes theorem?
- 39)What is meant by likelihood probability in naive Bayes theorem?
- 40)How we can deal with Missing values?What way to deal with missing value or null value.
- 41)What is NLTK?

- 42)What is Tokenization in NLP?
- 43)What is Stemming?
- 44)What is Lemmitization?
- 45)What is Corpus in NLP?
- 46)What is Spark framework?

Steps to run Scala program in windows with Spark framework?

- 1)First copy your scala file in folder spark in C drive/Program File/Big data/Spark
- 2)open command prompt(CMD) in same path with address bar.(Type in cmd in address bar of Spark Folder)
- 3)type Spark-shell command and press enter
- 4)to run scala file type below command and press enter

Example:

```
:load pali.scala  
: load sum.scala
```



Study material provided by: Vishwajeet Londhe

Join Community by clicking below links



Telegram Channel



https://t.me/SPPU_TE_BE_COMP

(for all engineering Resources)

SPPU Engineering & Technical
UPDATES
WhatsApp channel



WhatsApp Channel

(for all Engg & tech updates)



[https://whatsapp.com/channel/
0029ValjFriICVfpcV9HFc3b](https://whatsapp.com/channel/0029ValjFriICVfpcV9HFc3b)



Insta Page

(for all Engg & tech updates)



@SPPU_ENGINEERING_UPDATE

[https://www.instagram.com/
sppu_engineering_update](https://www.instagram.com/sppu_engineering_update)