

Q-92) _____ use for defining and containing geographical data which use by earth browsers.

- a) UML
- b) XML
- c) KML
- d) DML

Q-93) In which chart does the x-axis wrapped in a 360° round?

- a) Mekko chart
- b) Gauge chart
- c) Radar chart
- d) Combo chart

Q-94) What action do you expect to take place upon pressing Ctrl+E in QlikView?

- a) Reloading the script
- b) Opening script editor
- c) Open document properties
- d) Open report editors

Q-95) Which of the following is not the type of keep?

- a) Left
- b) Right
- c) Inner
- d) Outer

Q-96) _____ is used to create the mapping table that can be used for replacing field value and field names.

- a) Mapping table
- b) Mapping query
- c) Mapping function
- d) Mapping load

Q-96) What is considered to be the optimal data structure?

- a) Star schema
- b) Snowflake schema
- c) Database schema
- d) Synthetic schema

Q-97) Can Qlik View extract data from website?

- a) Yes
- b) No

Q-98) Which of the following is not a rotate option?

- a) Rotate right
- b) Rotate left
- c) Rotate out
- d) Transpose

Q-99) Ctrl + T is used to view _____

- a) Table view
- b) Dashboard
- c) Menu
- d) None

Q-100) Ctrl + R is used for _____

- a) Reload
- b) Rename
- c) Partial reload
- d) None

Q-101) Which of the following is not a type of load in Qlik View?

- a) Buffer load
- b) Residential Load
- c) Incremental Load
- d) Add load

Q-102) A binary load is one of the types of data loaded from QlikView file into the computer's _____.

- a) Drivers
- b) RAM
- c) ROM
- d) Register Memories

Q-103) Which of the following has only expressions but not dimensions.

- a) Gauge chart
- b) List Box
- c) Trellis chart
- d) Both (a) and (b)

Q-104) A _____ is a compressed table that shows multiple table fields

- a) List box
- b) Multi-box
- c) Statistic box
- d) None of the above

Q-105) _____ option is used while connecting to a database using ODBC and when only provides the 32-bit driver.

- a) Force 32 bit
- b) Force 43 bit
- c) Force 0 bit
- d) None

Q-106) ____ is where at least two tables comprises more than one normal section between them.

- a) Synthetic key
- b) Engineered Key
- c) Manufactured key
- d) None

Q-107) Ctrl + Shift + P is used for _____

- a) Print
- b) Print as image
- c) Print as PDF
- d) None

Q-108) A _____ is a concept that is created when the relationship between two tables can be established directly or through the third table.

- a) For loop
- b) While loop
- c) Circular loop
- d) None

Q-109) Which of the following is not the part of Qlik View Server environment?

- a) Qlik View web server
- b) Directory Service connector
- c) Qlik View Distribution Services
- d) Qlik View console management

Q-110) _____ is a platform that centrally manages all the other QlikView components.

- a) Qlik View Publisher.
- b) Qlik View server
- c) Qlik View Desktop
- d) Qlik Management Console

Q-111) Which of the following are document event triggers?

- a) On Open
- b) OnPostReload
- c) OnEnyselect
- d) OnPostRewduseData
- e) All of the above

Q-112) Which of the following are field event triggers?

- a) OnSelect
- b) OnChange
- c) OnLock
- d) OnUnlock
- e) All of the above

Q-113) A _____ maintains communication between client/end users.

- a) Qlik View Publisher.
- b) Qlik View server
- c) Qlik View Desktop
- d) Qlik View Directory

Q-114) P() is used for _____

- a) Possible values
- b) Excluded values
- c) Targeted values
- d) None of the above

Q-115) If we have data like '98as32r9' and we wish to remove all the alphabets. Which function we will use?

- a) Removealphabets()
- b) Purgechar()
- c) Removechar()
- d) None

Q-116) What is date#() function used for?

- a) Formatting the date
- b) Changing the date
- c) Identifying the date format
- d) None

Q-117) _____ treats the null values as missing values and does not allow linking of data.

- a) NullAsNull
- b) NullAsValue
- c) NullAsEmpty
- d) None

Q-118) How many types of triggers are available in Qlik View?

- a) 1
- b) 2
- c) 3
- d) 4

Q-119) Ctrl + B is used for _____

- a) Adding schema
- b) Adding Table
- c) Adding record
- d) Adding Bookmark

Q-120) Which of the following are book marks in Qlik View?

- a) Document Bookmarks.
- b) User Bookmarks.
- c) Private Server Bookmarks.
- d) Shared Server Bookmarks.

Q-121) Using what we can show multiple expression in table box?

- a) ||
- b) & &
- c) !!
- d) ==

Q-122) _____ allows customers to access and manipulate data files.

- a) Qlik View Publisher.
- b) Qlik View User Access
- c) Qlik View Desktop
- d) Qlik View Directory

Q-123) What is book mark?

- a) A bookmark can select the data in all the states mentioned in the QlikView document.

- b) A bookmark can capture the selections in all the states mentioned in the QlikView document.
- c) A bookmark can load the data in all the states mentioned in the QlikView document.
- d) None

Q-124) _____ system variable used to get the number of statements that have caused errors during script execution.

- a) ScriptErrorCount
- b) ScriptError
- c) ScriptCount
- d) SEC

Q-125) E() is used for _____

- a) Possible values
- b) Excluded values
- c) Targeted values
- d) None of the above

Q-126) In layout tab for sheet objects top, normal and bottom are denoted by what?

- a) 0,1,-1
- b) 1,-1,0
- c) 0,-1,1
- d) -1,0,1

Q-127) _____ helps users to engage in threaded discussions with content.

- a) Qlik View chat

- b) Qlik View FAQs
- c) Qlik View Server
- d) Qlik View Annotation

Q-128) _____ is used by developers to use to create the QlikView data model and GUI layout for application to use at the frontend.

- a) Qlik View Publisher.
- b) Qlik View server
- c) Qlik View Desktop
- d) Qlik View Directory

Q-129) What are various types of tables in Qlik View?

- a) Cross tables
- b) Pivot tables
- c) Straight tables
- d) Mapping table
- e) All of the above

Q-130) Which chart has the following look types: Speedometer, Reflected Tube, Traffic light, Test Tube

- a) Gauge chart
- b) Test tube chart
- c) Radar chart
- d) Funnel chart

Q-131) Which of the following are triggers?

- a) Document Trigger
- b) Sheet trigger
- c) Server Trigger
- d) All of the above

Q-132) Which of the following are document triggers?

- a) Document event trigger
- b) Field Event trigger
- c) Variable Event trigger
- d) All of the above

Q-133) Ctrl + E is used to _____

- a) Edit the page
- b) Edit the script
- c) Edit the charts
- d) None

Q-134) We can find total number of records by using which of the following?

- a) QvdNumberOfRecords ('path with filename')
- b) QvsNumberOfRecords ('path with filename')
- c) QvPNumberOfRecords ('path with filename')
- d) None

Q-135) How many types of bookmark is available in Qlik View?

- a) 1
- b) 2
- c) 3
- d) 4

Q-136) _____ enables linking of data that are null.

- a) NullAsValue
- b) NullAsNull
- c) NullAsEmpty
- d) None

Q-137) By which function does the second half of the table continues besides the first half rather than below it?

- a) Wrap
- b) Rotate
- c) Column
- d) Fill

Q-138) What is considered as identifier in set analysis?

- a) \$
- b) &
- c) *
- d) !

Q-139) Ctrl +Shift + R is used for _____

- a) Partial reload
- b) Reload script
- c) Reload charts
- d) none

Q-140) The _____ is a way of sorting and arranging data into different sets in sheet objects like charts or tables.

- a) Set analysis
- b) Squat analysis
- c) Cross analysis
- d) None

Q-141) Full form of XML is _____.

- a) Extensive Markup language
- b) Extra Markup language
- c) Extensible Markup language
- d) None

Q-142) _____ function these function we are used for Update and delete records updating in QVD. these also we are using where clause.

- a) Exist
- b) Not Exist
- c) Exit
- d) None

Q-143) From which one of these menus will you get an option for Alerts?

- a) Layout
- b) Views
- c) Tool
- d) Object

Q-144) _____ is handled by IT professionals in Qlik View.

- a) Qlik View Publisher.
- b) Qlik View server
- c) Qlik View Desktop
- d) Qlik Management Console

Q-145) _____ function is used commonly in where clause is the values need to load whatever we have already in QVD those records need to load that time

- a) Exist
- b) Not Exist
- c) Exit
- d) None

Q-146) How can we drop some fields from the memory during script execution?

- a) Drop
- b) Delete
- c) Exit
- d) None

Q-147) Which of the following is not used to separate data fields from a delimited files?

- a) Comma
- b) Tab
- c) Semicolon
- d) Full stop

Q-149) Can we store a sound file in a QlikView document?

- a) Yes
- b) No\

Q-150) We can store a sound file in Qlik View Document using which prefix?

- a) Bundle
- b) Sound
- c) Voice
- d) None

Q-151) _____ statement is used to load data from databases.

- a) Load
- b) Select
- c) Match
- d) All

Q-152) _____ expansion is used to expand the contents of one cell into several cells in the table.

- a) Content Cell

- b) Context Cell
- c) Call Cell
- d) None

Q-153) Applications created in the desktop version are hosted on _____ so that users can access and use them anywhere.

- a) Qlik View Desktop
- b) Qlik View publisher
- c) Qlik View Server
- d) Qlik View Console

Q-154) The _____ mode uses the internal web browser in QlikView to display the document layout as an AJAX page.

- a) Web browser
- b) Web view
- c) Web scrapper
- d) None

Q-155) What does *si* in text search means?

- a) It searches word with Si as start
- b) It searches word with Si as end
- c) It searches any string that contains si
- d) None

Q-156) While loading two tables cannot intersect with each other. How to find out in QV?

- a) Join

- b) Outer Join
- c) Concatenate
- d) NoConcatenate

Q-157) Which option is used to create macros?

- a) Macros module
- b) Micros module
- c) Edit module
- d) All of the above

Q-158) _____ statement is used to load data from all files.

- a) Load
- b) Select
- c) Match
- d) All

Q-159) ETL and OLAP is supported in which of the following?

- a) Qlik View.
- b) Qlik Sense

Q-160) Which of the following has hidden script?

- a) Qlik View.
- b) Qlik Sense

Q-161) Which of the following supports multi cloud platform?

- a) Qlik View.
- b) Qlik Sense

Q-162) Which button helps you to enter in search mode?

- a) F3
- b) F4
- c) F5
- d) F6

Q-163) The _____ function to a Load or Select (SQL) statement is used for linking discrete numeric values to one or more numeric intervals.

- a) Match
- b) Internalmatch
- c) NoMatch
- d) None

Q-164) ctrl + G is used for _____

- a) Show/hide ruler
- b) Show/hide grid
- c) Show/hide script
- d) none

Q-165) The _____ option is used to mark and delete the data that is not required or that is jumbled and not useful.

- a) Dustbin
- b) Garbage

- c) Throw
- d) All of the above

Q-166) Which of the following is not the part of Qlik View architecture?

- a) Qlik View Publisher.
- b) Qlik View server
- c) Qlik View Desktop
- d) Qlik View Directory

Q-167) What are the two ways to insert Load statements into script editor?

- a) Inline
- b) File
- c) Both a and b
- d) None

Q-168) ____ is where at least two tables comprises more than one normal section between them.

- a) Synthetic key
- b) Engineered Key
- c) Manufactured key
- d) None

Q-169) How many types of rotate option are available in Qlik View?

- a) 1
- b) 2

c) 3

d) 4

Q-170) A _____ shows all the values of single table fields in the form of a list.

- a) List box
- b) Multi-box
- c) Statistic box
- d) None of the above



SUBJECTIVE QUESTION.

- Q1) What are the difference in the features of Power BI, Tableau, and Qlik View?**
- Q2) Explain the components of Qlik View. Explain its features?**
- Q3) What is difference between Qlik view and Qlik Sense?**
- Q4) What are different types of charts which you can use in Qlik View?**
- Q5) Name some tiers of Qlik View and explain them?**
- Q6) Explain snapshots in qv**
- Q7) Explain .QVD files in Qlik View?**
- Q8) Explain CAL in qlik view?**
- Q9) Explain difference between Join and keep.**
- Q10) Explain types of join in Qlik view.**
- Q11) Explain trellis chart of qlik view**
- Q12) Explain difference between straight table and pivot table**
- Q13) Explain snowflake scheme**
- Q14) Explain various statistical functions used in qlik view?**
- Q15) What are various date and time functions in qlik view?**
- Q16) Explain Sheets options in QlikView?**
- Q17) What is set analysis in qlik view?**
- Q18) What is the QlikView server and Publisher?**
- Q19) Name different data bases that can be used with Qlik view?**
- Q20) Explain crosstable feature of qlik view?**
- Q21) Explain scripting in qlik view?**
- Q22) Explain match() function of QV?**
- Q23) What do you know about the report interface of QV?**
- Q24) Significance of rank() function of QV?**
- Q25) What is QIX Engine?**

Q26) Explain data transformation feature of QV?

Q27) Describe the data model view of QV?

Q28) What is star schema in QV?

Q29) Discuss master calendar option of qv?

Q30) What is fuzzy search in qv?



R Interview Questions



Instructions to solve the Questions:

- 1. Explanation for each and every solution must be given in brief.**
 - 2. Screenshot for each and every step must be provided for questions involving practical implementation.**
 - 3. For theory/conceptual questions. Solutions having at least 1 practical implementation will be favoured.**
-

-
- 1) List out some of the function that R provides?
- 2) Explain how you can start the R commander GUI?
- 3) In R how you can import Data?
- 4) Explain how R commands are written?
- 5) How can you save your data in R?
- 6) Mention how you can produce co-relations and covariances?
- 7) Explain what is t-tests in R?
- 8) Explain what is With () and By () function in R is used for?
- 9) What are the data structures in R that is used to perform statistical analyses and create graphs?
- 10) Explain general format of Matrices in R?
- 11) In R how missing values are represented?
- 12) Explain what is transpose?
- 13) Explain how data is aggregated in R?
- 14) What is the function used for adding datasets in R?
- 15) What is the use of subset() function and sample() function in R ?

- 16)** Explain how you can create a table in R without external file?
- 17)** What are the different components of grammar of graphics?
- 18)** What is Rmarkdown? What is the use of it?
- 19)** How do you install a package in R?
- 20)** What are the steps to build and evaluate a linear regression model in R?
- 21)** Name some packages in R, which can be used for data imputation?
- 22)** Explain about confusion matrix in R?
- 23)** Explain initialize() function in R?
- 24)** How can we find the mean of one column with respect to another?
- 25)** What is a Random Walk model?
- 26)** What is a White Noise model?
- 27)** Give any five features of R.
- 28)** Differentiate between R and Python in terms of functionality?
- 29)** Explain RStudio.
- 30)** What are the advantages and disadvantages of R?
- 31)** What is the purpose behind R and Hadoop integration?
- 32)** What will be the output of the expression all(NA==NA)?
- 33)** What is the difference b/w sample() and subset() in R?
- 34)** Why do we use the command - install.packages(file.choose(), repos=NULL)?
- 35)** Give the command to create a histogram and to remove a vector from the R workspace?
- 36)** Differentiate b/w "%%" and "%/%".
- 37)** Why do we use apply() function in R?
- 38)** Differentiate between library() and require() functions.
- 39)** What is the t-test() in R?
- 40)** What is the use of with() and by() functions in R?
- 41)** Differentiate b/w lapply and sapply.
- 42)** Explain aggregate() function.
- 43)** Explain the doBy package?
- 44)** Explain fitdistr() function?
- 45)** What are GGobi and iPlots?
- 46)** Explain the lattice package.
- 47)** Explain anova() function.
- 48)** Explain cv.lm() and stepAIC() function
- 49)** Explain leaps() function

- 50)** Explain relaimpo and robust package
- 51)** Give full form of MANOVA and what is the use of it.
- 52)** Explain mashapiro.test() and barlett.test().
- 53)** Explain the use of the forecast package.
- 54)** Differentiate between qda() and lda() function.
- 55)** Explain the auto.arima() and principal() function.
- 56)** Explain FactoMineR
- 57)** What is the full form of SEM and CFA?
- 58)** Define cluster.stats() and pvclust() function().
- 59)** Define MATLAB and party packages.
- 60)** Explain S3 and S4 systems.
- 61)** Explain Chi-Square Test
- 62)** Explain Random Forest.
- 63)** Explain Time Series Analysis.
- 64)** Explain Pie chart in R.
- 65)** Explain Histogram.
- 66)** What is the main difference between an Array and a matrix?
- 67)** Which data object in R is used to store and process categorical data?
- 68)** What is the main difference between an Array and a matrix?
- 69)** Which data object in R is used to store and process categorical data?
- 70)** How can you load and use csv file in R?
- 71)** How do you get the name of the current working directory in R?
- 72)** What is R Base package?
- 73)** How R is used in logistic regression?
- 74)** How do you access the element in the 2nd column and 4th row of a matrix named M?
- 75)** What is recycling of elements in a vector?
- 76)** Give an example.What are different ways to call a function in R?
- 77)** What is lazy function evaluation in R?
- 78)** How do you install a package in R?
- 79)** Name a R packages which is used to read XML files.
- 80)** Can we update and delete any of the elements in a list?
- 81)** Give the general expression to create a matrix in R.
- 82)** which function is used to create a boxplot graph in R?
- 83)** In doing time series analysis, what does frequency = 6 means in the ts() function?

- 84)** What is reshaping of data in R?
- 85)** What is the output of runif(4)?
- 86)** How to get a list of all the packages installed in R ?
- 87)** What is expected from running the command - strsplit(x,"e")?
- 88)** Give a R script to extract all the unique words in uppercase from the string - "The quick brown fox jumps over the lazy dog".
- 89)** Vector v is c(1,2,3,4) and list x is list(5:8), what is the output of v*x[1]?Vector v is c(1,2,3,4) and list x is list(5:8), what is the output of v*x[[1]]?
- 90)** What does unlist() do?
- 91)** Give the R expression to get 26 or less heads from a 51 tosses of a coin using pbinom.
- 92)** X is the vector c(5,9.2,3,8.51,NA), What is the output of mean(x)?
- 93)** How do you convert the data in a JSON file to a data frame?
- 94)** Give a function in R that replaces all missing values of a vector x with the sum of elements of that vector?
- 95)** What is the use of apply() in R?
- 96)** Is an array a matrix or a matrix an array?
- 97)** How to find the help page on missing values?
- 98)** How do you get the standard deviation for a vector x?
- 99)** How do you set the path for current working directory in R?
- 100)** What is the difference between "%%" and "%/%"?
- 101)** What does col.max(x) do?
- 102)** Give the command to create a histogram.
- 103)** How do you remove a vector from the R workspace?
- 104)** List the data sets available in package "MASS"
- 105)** List the data sets available in all available packages.
- 106)** What is the use of the command - install.packages(file.choose(), repos=NULL)?
- 107)** Give the command to check if the element 15 is present in vector x.
- 108)** Give the syntax for creating scatterplot matrices.
- 109)** What is the difference between subset() function and sample() function in R?
- 110)** How do you check if "m" is a matrix data object in R?
- 111)** What is the output for the below expression all(NA==NA)?
- 112)** How to obtain the transpose of a matrix in R?
- 113)** What is the use of "next" statement in R?
- 114)** Which function in R language is used to find out whether the means of 2 groups are equal to each other or not?
- 115)** What is the best way to communicate the results of data analysis using R language?

- 116)** What are with () and BY () functions used for?
- 117)** dplyr package is used to speed up data frame management code. Which package can be integrated with dplyr for large fast tables?
- 118)** In base graphics system, which function is used to add elements to a plot?
- 119)** What are the different type of sorting algorithms available in R language?
- 120)** What is the command used to store R objects in a file?
- 121)** What is the best way to use Hadoop and R together for analysis?
- 122)** What will be the output of log (-5.8) when executed on R console?
- 123)** How is a Data object represented internally in R language?
- 124)** Which package in R supports the exploratory analysis of genomic data?
- 125)** What is the difference between data frame and a matrix in R?
- 126)** How can you add datasets in R?
- 127)** What are factor variable in R language?
- 128)** What is the memory limit in R?
- 129)** What are the data types in R on which binary operators can be applied?
- 130)** How do you create log linear models in R language
- 131)** What will be the class of the resulting vector if you concatenate a number and NA?
- 132)** What is meant by K-nearest neighbour?
- 133)** What will be the class of the resulting vector if you concatenate a number and a character?
- 134)** If you want to know all the values in c (1, 3, 5, 7, 10) that are not in c (1, 5, 10, 12, 14).
Which in-built function in R can be used to do this? Also, how this can be achieved without using the in-built function.
- 135)** How can you debug and test R programming code?
- 136)** What will be the class of the resulting vector if you concatenate a number and a logical?
- 137)** Write a function in R language to replace the missing value in a vector with the mean of that vector.
- 138)** What happens if the application object is not able to handle an event?
- 139)** Differentiate between seq (6) and seq_along (6)
- 140)** How do you write R commands?
- 141)** How can you verify if a given object "X" is a matrix data object?
- 142)** What do you understand by element recycling in R?
- 143)** How can you verify if a given object "X" is a matrix data object?
- 144)** How will you measure the probability of a binary response variable in R language?
- 145)** What is the use of sample and subset functions in R programming language?
- 146)** There is a function fn(a, b, c, d, e) a + b * c - d / e. Write the code to call fn on the vector c(1,2,3,4,5) such that the output is same as fn(1,2,3,4,5).
- 147)** How can you resample statistical tests in R language?
- 148)** What is the purpose of using Next statement in R language?
- 149)** How will you create scatterplot matrices in R language?
- 150)** How will you check if an element 25 is present in a vector?
- 151)** What is the difference between library() and require() functions in R language?
- 152)** What are the rules to define a variable name in R programming language?
- 153)** What do you understand by a workspace in R programming language?
- 154)** Which function helps you perform sorting in R language?
- 155)** How will you list all the data sets available in all R packages?
- 156)** Which function is used to create a histogram visualisation in R programming language?
- 157)** Write the syntax to set the path for current working directory in R environment.
- 158)** How will you drop variables using indices in a data frame
- 159)** What will be the output of runif (7)?
- 160)** What is the difference between rnorm and runif functions ?
- 161)** What will be the output on executing the following R programming code –

```
mat<-matrix(rep(c(TRUE,FALSE),8),nrow=4)
```

```
sum(mat)
```

- 162) How will you combine multiple different string like “Data”, “Science”, “in”, “R”, “Programming” as a single string “Data_Science_in_R_Programmming” ?
- 163) Write a function to extract the first name from the string “Mr. Tom White”.
- 164) Can you tell if the equation given below is linear or not?
- 165) What will be the output of the following R programming code?

```
var2<- c("I","Love,"ProjectPro")
```

```
var2
```

- 166) I have a string "contact@dezyre.com". Which string function can be used to split the string into two different strings “contact@dezyre” and “com”
- 167) What is R Base package?
- 168) How will you merge two dataframes in R programming language?
- 169) Explain the usage of which() function in R language
- 170) How will you convert a factor variable to numeric in R language ?
- 171) What is power analysis ?
- 172) Explain the usage of abline() function
- 173) What is the usage of lattice package in R ?
- 174) When is it appropriate to use the “next” statement in R?
- 175) How do you assign a variable in R?
- 176) What are the different data types/objects in R
- 177) How R commands are written?
- 178) What is the use of subset() and sample() function in R?
- 179) How you can produce co-relations and covariances?
- 180) Which method is used for exporting the data in R?
- 181) Which packages are used for exporting of data?
- 182) How impossible values are represented in R?
- 183) Which command is used for storing R object into a file?
- 184) Which command is used for restoring R object from a file?
- 185) What is the use of coin package in R?
- 186) What is the use of tapply?
- 187) What happens when the application object does not handle an event?
- 188) Explain app specific objects which store the app contents
- 189) Explain the purpose of using UIWindow object?
- 190) How to create axes in the graph?
- 191) Why vcd package is used?
- 192) What is iPlots?
- 193) What is the use of lattice package?

- 194)** What is fitdistr() function?
 - 195)** What is the use of sink() function?
 - 196)** On which type of data binary operators are worked?
 - 197)** What is the use of doBY package?
 - 198)** Define loglm() function
 - 199)** How to create scatterplot matrices?
 - 200)** What is the use of diagnostic plots?
-



Statistics Interview Questions

Guidelines

- No Plagiarism will be entertained in the subjective questions
 - The main aim of all the questions where the answer is numeric value is not the answer, but the approach you followed to get that answer so make sure to explain every step in a clean way.
 - Make sure visuals are easily interpretable and less cluttered so follow the rules of data visualization.
 - Give real-world examples as much as you can in the answers. Good Luck!
-

Q-1 Explain Correlation? With real-world examples

Q-2 Briefly Explain why correlation does not imply causation. With real-world examples

Q-3 What is Covariance. And what is the main difference between correlation & Covariance?

Q-4 What are the 2 main Fundamental plots to catch Correlation?

Q-5 What is a z-score? EXPLAIN intuitively.

Q-6 Explain Normal distribution with its properties. With real-world examples

Q-7 How is a Normal distribution different from a Standard Normal Distribution? Is there a way to convert it into one another? If Yes, Explain briefly that method.

Q-8 What are Outliers? Explain with a mathematical term.

Q-9 What is a left-skewed and right-skewed distribution. Explain Briefly

Q-10 What is a Symmetric distribution? Give all examples of that distribution with some visuals.

Q-11 What is a relationship between mean, the median in Normal Distribution?

Q-12 What do these terms Bell Curve and Gaussian Distribution mean?

Q-13 How is a box-plot different from a scatterplot?

Q-14 What is the difference between Descriptive and Inferential statistics?

Q-15 What is a population and sample in inferential statistics. Explain briefly with good examples.

Q-16 Let A and B be events on the same sample space, with $P(A) = 0.3$ and $P(B) = 0.8$. Can these two events be disjoint?

Q-17 Can we use a z-table for a skewed distribution? Explain

Q-18 Which statistic is most sensitive to outlier – Mean, Median or mode?

Q-19 What is a percentile?

Q-20 Explain the effect of Shifting linear Transformation on mean, median, S.D. & IQR of data?

Q-21 Explain the effect of Scaling linear Transformation on mean, median, S.D. & IQR of data?

Q-22 What's Density Curve. Give a diagram also.

Q-23 Briefly explain the 'Empirical Rule' in statistics. Explaining with a diagram is mandatory, preferably with real-world data examples.

Q-24 What is the mean and S.D. of a Standard normal distribution.

Q-25 Lifespan of African Lizards is normally distributed with mean = 3.1yrs and S.D. = 0.6yrs . Find the Prob. Of Lizard living between 2.5 yrs to 4.3 yrs. Explain visually with Normal distribution.

Q-26 Consider a Normal distribution of student's heights with mean = 15cm and S.D. = 20cm. Rohan's height = 161.4cm.

What Prob. Of students have height less than Rohan's? Explain visually with Normal distribution

Q-27 Consider a Normal distribution of student's heights with mean = 15cm and S.D. = 20cm. Rohan's height = 161.4cm.

What Prob. Of students have height more than Rohan's? Explain visually with Normal distribution

Q-28 What is the Prob of a student having the same height? Give an approximate value (Hint: No calc. is needed here)

Q-29 A set of laptop prices are normally dist. Mean = \$750 and S.D. = \$60. What is the prob? Of lap prices between \$624 and \$768. Explain visually with Normal distribution.

Q-30 Consider Normal dist. Of pulse rates. Mean = 80, S.D. = 9. Additional testing is to be done for the top 40%. Find the minimum pulse rate among those top 40%? Explain visually with Normal distribution

Q-31 Normal dist of avg. wait time for various roads in Lucknow. Mean = 185 secs and S.D. = 11 secs. Anand wants to take the road that is in the bottom 10% of Lucknow. What is the max wait time in that bottom 10%? Explain visually with Normal distribution

Q-32 State the formula for the correlation coefficient

Q-33 Explain each part of the formula you mentioned above, preferably with an example.

Q-34 Ram has 2 children and 1 of them is a girl. What is the probability that the other child is also a girl?

You can assume that there is an equal number of males and females in the world. Define sample spaces properly.

Q-35 A fair six-sided die is rolled twice. What is the probability of getting 2 on the first roll and not getting 4 on the second roll?

Q-36

$$P(A \cup B \cup C) = P(A \cap C^c) + P(C) + P(B \cap A^c \cap C^c)$$

Is it True or False? Explain your answer.

Q-37 Consider a 4-sided die and roll it twice. What is the probability that the number on the first roll is strictly higher than the number on the second roll?

Define sample space properly.

Q-38 Which of the following can't be a Probability event:

- 0.1
- -0.04
- 0.5
- 0.99999
- 1.0001
- 1/89
- 0

Explain why's with your answer.

Q-39 Urvin randomly picks 4 cards from a deck of 52-cards and places them back into the deck (Any set of 4 cards is equally likely). Then, Amit randomly chooses 8 cards out of the same deck (Any set of 8 cards is equally likely). Assume that the choice of 4 cards by Urvin and the choice of 8 cards by Amit are independent. What is the probability that all 4 cards chosen by Urvin are in the set of 8 cards chosen by Amit? Explain

Context: A player is randomly dealt a sequence of 13 cards from a deck of 52-cards. All sequences of 13 cards are equally likely. In an equivalent model, the cards are chosen and dealt one at a time. When choosing a card, the dealer is equally likely to pick any of the cards that remain in the deck.

Q- 40 If you dealt 13 cards, what is the probability that the 13th card is a King?

Q-41 A fair six-sided die is rolled 6 times. What is the probability of getting all outcomes as unique?

Q-42 A group of 60 students is randomly split into 3 classes of equal size. All partitions are equally likely. A and B are two students belonging to that group. What is the probability that A and B will end up in the same class?

Q-43 We have two coins, A and B. For each toss of coin A, the probability of getting head is $1/2$ and for each toss of coin B, the probability of getting heads is $1/3$. All tosses of the same coin are independent. We select a coin at random. and toss it till we get a head. The probability of selecting coin A is $\frac{1}{4}$ and coin B is $3/4$. What is the expected number of tosses to get the first heads?

Q-44 How would u calculate range & interquartile range.

Q-45 Differentiate between quantitative and qualitative data.

Q-46 Explain with examples -- Categorical ordinal and nominal variables

Q-47 Explain with examples – Numerical Continuous and Discrete data.

Q-48 What is the five-number summary and how is it useful?

Q-49 Suppose a life insurance company sells a \$240,000 one-year term life insurance policy to a 25-year-old female for \$210. The probability that the female survives the year is .999592. Find the expected value of this policy for the insurance company.

Q-50 When can a probability event be independent of itself?

Q-51 Name and explain any 3 outlier detection techniques.

Q-52 Name and explain any 3 outlier handling techniques

Q-53 What would be your approach if I ask you the mean length of all road trips in India? (With a confidence of 95%)

Q-54 Can you relate Standard deviation with Standard error. Explain

Q-55 What are the various effects/trade-offs of changing the width of Confidence Intervals?

Q-56 On increasing Margin of error, what happens to C.I.?

Q-57 What is the difference between 95% confidence and 99% confidence. Can you give a scenario for both?

Q-58 Explain the Degree of freedom

Q-59 What do you think will happen if DF (deg. Of freedom) is large?

Q-60 What is Statistical power?

Q-61 What are a z-test and t-test? Explain briefly

Q-62 When to use which distribution/test? (z & t-test)

Q-63 What are null and alternate hypotheses?

Q-64 Give examples for the above question for one-sided and 2-sided tests.

Q-65 What is the p-value? Explain briefly with example and visual

Q-66 How would you conduct a hypothesis test? Briefly tell every step.

Q-67 Calculate the p-value using any of your favourite tool (python, excel etc) and describe it here.

Q-68 What do we mean by – making a decision based on comparing p-value with significance level?

Q-69 If the null hypothesis consists of a constant value only, what type of test is it you think? Explain in your best way.

Q-70 Do you think the critical values of 95% one tail and 90% two tail are same. If yes, why?

Q-71 Cross-fertilizing a red and a white flower produces red flowers 25% of the time. Now we cross-fertilize five pairs of red and white flowers and produce

five offspring. What is the probability that there are no red flower plants in the five offspring?

Q-72 A roulette wheel has 38 slots – 18 red, 18 black, and 2 green. You play five games and always bet on red slots. How many games can you expect to win?

Q-73 A roulette wheel has 38 slots, 18 are red, 18 are black, and 2 are green. You play five games and always bet on red. What is the probability that you win all the 5 games?

Q-74 Some test scores follow a normal distribution with a mean of 18 and a standard deviation of 6. What proportion of test takers have scored between 18 and 24?

Q-75 A jar contains 4 marbles. 3 Red & 1 white. Two marbles are drawn with replacement after each draw. What is the probability that the same colour marble is drawn twice?

Q-76 Which of the following events is most likely? Give why

- A) At least one 6, when 6 dice are rolled
- B) At least 2 sixes when 12 dice are rolled
- C) At least 3 sixes when 18 dice are rolled
- D) All the above have same probability

Q-77 What is the main difference between Overfitting and underfitting?

Q-78 What is selection bias in statistics?

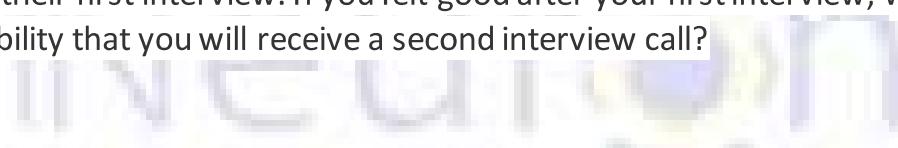
Q-79 What is a Sampling Distribution?

Q-80 What is the Law of Large Numbers?

Q-81 What is the Central Limit Theorem? And what's the use?

Q-82 Briefly explain type-I and type-II Error? Give diagrams also.

Q-83 Suppose you were interviewed for a technical role. 50% of the people who sat for the first interview received the call for second interview. 95% of the people who got a call for second interview felt good about their first interview. 75% of people who did not receive a second call, also felt good about their first interview. If you felt good after your first interview, what is the probability that you will receive a second interview call?



Q-84 A coin of diameter 1-inches is thrown on a table covered with a grid of lines each two inches apart. What is the probability that the coin lands inside a square without touching any of the lines of the grid? You can assume that the person throwing has no skill in throwing the coin and is throwing it randomly.

You can assume that the person throwing has no skill in throwing the coin and is throwing it randomly.

Q-85 Consider the following probability density function: What is the probability for $X \leq 6$ i.e. $P(X \leq 6)$

$$f(x) = \frac{1}{8} e^{-x/8} \text{ for } x \geq 0$$

What is the probability for $X \leq 6$ i.e. $P(X \leq 6)$

Q-86 In a class of 30 students, approximately what is the probability that two of the students have their birthday on the same day (defined by same day and month) (assuming it's not a leap year)?

For example – Students with birthday 3rd Jan 1993 and 3rd Jan 1994 would be a favourable event.

Q-87 Assume you sell sandwiches. 70% people choose Veg, and the rest choose non-veg. What is the probability of selling 2 veg sandwiches to the next 3 customers?

Q-88 Anand is playing a lottery game where he must pick 2 numbers from 0 to 9 followed by an English alphabet (from 26-letters). He may choose the same number both times.

If his ticket matches the 2 numbers and 1 letter drawn in order, he wins the grand prize and receives \$10405. If just his letter matches but one or both of the numbers do not match, he wins \$100. Under any other circumstance, he wins nothing. The game costs him \$5 to play. Suppose he has chosen 04R to play.

What is the expected net profit from playing this ticket?

Context: HIV is still a very scary disease to even get tested for. The US military tests its recruits for HIV when they are recruited. They are tested on three rounds of Elisa(an HIV test) before they are termed to be positive.

The prior probability of anyone having HIV is 0.00148. The true positive rate for Elisa is 93% and the true negative rate is 99%.

Q-89 What is the probability that a recruit has HIV, given he tested positive on first Elisa test? The prior probability of anyone having HIV is 0.00148. The true positive rate for Elisa is 93% and the true negative rate is 99%.

Q-90 What is the probability of having HIV, given he tested positive on Elisa the second time as well.

The prior probability of anyone having HIV is 0.00148. The true positive rate for Elisa is 93% and the true negative rate is 99%.

Q-91 Suppose you're playing a game in which we toss a fair coin multiple times. You have already lost thrice where you guessed heads but a tails appeared. Which of the below statements would be correct in this case?

- A) You should guess heads again since the tails has already occurred thrice and its more likely for heads to occur now
- B) You should say tails because guessing heads is not making you win
- C) You have the same probability of winning in guessing either, hence whatever you guess there is just a 50-50 chance of winning or losing
- D) None of these

Q-92 What is Bayesian?

Q-93 What is Frequentist?

Q-94 What is Likelihood?

Q-95 The inference using the frequentist approach will always yield the same result as the Bayesian approach. **Yes/no with Reason**

Q-96 What is Quantile? Is there a difference between quantile and quartile?

Q-97 What is the difference between skewness and kurtosis? Give answer explaining visually

Q-98 What is Moment?

Q-99 State the binomial probability formula.

Q-100 Hospital records show that 75% of patients suffering from a disease die due to that disease. What is the probability that 4 out of the 6 randomly selected patients recover?

Q-101 Give and explain the types of Statistical Sampling techniques.

Q-102 What is the main difference between Data-Mining and Data Profiling?

Q-103 How would you deal with a multi-source data problems?

Q-104 What are types of Missing values.

Q-105 For each type stated above explain what it is with a real world example.

Q-106 Can a machine learning algorithm be used for imputing missing values. If yes, then explain a few methods/algos.

Q-107 What is KPI. Give a real world example

Q-108 What is Design of experiment? (w.r.t Statistics)

Q-109 What is the 80/20 rule of Businesses?

Q-110 What is Hot-Deck and Cold-Deck Imputation?

Q-111 What is the probability of throwing two fair dice when the sum is 5 and 8?

Q-112 State the case where median is better than mean for imputing missing data? Give a real world example.

Q-113 What is six-sigma in statistics?

Q-114 Given a left-skewed distribution that has a median of 60, what conclusions can we draw about the mean and the mode of the data?

Q-115 Discuss some scenarios where you will keep the outliers in data?

Q-116 If there is a 30% prob. Of seeing a car in any 20-min interval . State the prob. Of seeing atleast one car in an hour?

Q-117 Are Symmetric distributions always unimodal?

Q-118 What is Bessel's Correction?

Q-119 Differentiate between Univariate, Bivariate and Multivariate analysis?

Q-120 What types of variables are used for Pearson's correlation coefficient?

Q-121 In a scatter diagram, what is the line that is drawn above or below the regression line called?

Q-122 In an observation, there is a high correlation between the time a person studies and the amount of marks he gets. What can be inferred from this?

Q-123 What type of data does not have a log-normal distribution or a Gaussian distribution?

Q-124 What would you do if your data is not normal? Discuss your approach in brief here.

Q-125 Any 5 types of data transformations techniques to turn features to normal features.

Q-126 The students of a particular class were given two tests for evaluation. Twenty-five percent of the class cleared both the tests and forty-five percent of the students were able to clear the first test.

Calculate the percentage of students who passed the second test given that they were also able to pass the first test.

Q-127 While it is said that the probabilities of having a boy or a girl are the same, let's assume that the actual probability of having a boy is slightly higher at 0.51. Suppose a couple plans to have 3 children. What is the probability that exactly 2 of them will be boys?

Q-128 Heights of 10 year-olds, regardless of gender, closely follow a normal distribution with mean 55 inches and standard deviation 6 inches. Which of the following is true?

- A) We would expect more number of 10 year-olds to be shorter than 55 inches than the number of them who are taller than 55 inches
- B) Roughly 95% of 10 year-olds are between 37 and 73 inches tall
- C) A 10-year-old who is 65 inches tall would be considered more unusual than a 10-year-old who is 45 inches tall
- D) None of these

Q-129 About 30% of human twins are identical, and the rest are fraternal. Identical twins are necessarily the same sex, half are males and the other half are females. One-quarter of fraternal twins are both males, one-quarter both female, and one-half are mixed: one male, one female. You have just become a parent of twins and are told they are both girls. Given this information, what is the probability that they are identical?

Q-130 Rob has fever and the doctor suspects it to be typhoid. To be sure, the doctor wants to conduct the test. The test results positive when the patient actually has typhoid 80% of the time. The test gives positive when the patient does not have typhoid 10% of the time. If 1% of the population has typhoid, what is the probability that Rob has typhoid provided he tested positive?

Q-131 Arun is having two coins in his hand. Out of the two coins, one is a real coin and the second one is a faulty one with Tails on both sides. He blindfolds himself to choose a random coin and tosses it in the air. The coin falls down with Tails facing upwards. What is the probability that this tail is shown by the faulty coin?

Q-132 A fly has a life between 4-6 days. What is the probability that the fly will die at exactly 5 days ?

Q-133 What type of graph would you use for trend analysis. Give a visual also.

Q-134 Explain Survivorship bias.

Q-135 What are confounding variables?

Q-136 What is a point estimate (in inferential stats)?

Q-137 Explain the Experiment Design types in statistics. Give examples if you can.

Q-138 What is the Conditional Probability? Explain with example

Q-139 What is Bayes theorem, give formula also, explaining it. Explain with example.

Q-140 Difference between Binomial and Bernoulli distribution.

Q-141 What is a Poisson Distribution?

Q-142 Differentiate between Random Variables & Geometric Random variables.

Q-143 Give the formula for computing confidence interval for mean for a sample.

Q-144 Give the formula for computing confidence interval for proportion for a sample.

Q-145 Discuss the Assumption of Confidence interval in brief.

Q-146 State the assumption of Central Limit theorem.

Q-147 What is bootstrapping ? When is it useful in statistics?

Q-148 What is Simple Linear regression? How is it Different from Multiple Linear regression? Explain using real world examples.

Q-149 What is regression analysis? State the four most fundamental assumptions of regression? (you can tell more if you know)

Q-150 Explain the above assumptions in detail with examples.

Q-151 What is an A/B Test?

Q-152 Explain the steps involved in conducting an A/B Test?

