JDA Interview Questions

Courses 1-3

1. What is Conditional Formatting? How can it be used?

Answer:A conditional format changes the appearance of cells based on conditions that you specify. If the conditions are true, the cell range is formatted; if the conditions are false, the cell range is not formatted.

1. What is the difference between COUNT, COUNTA, COUNTBLANK, and COUNTIF in Excel?

Answer:

* COUNT function returns the count of numeric cells in a range.
* COUNTA function counts the non-blank cells in a range.
* COUNTBLANK function gives the count of blank cells in a range.
* COUNTIF function returns the count of values by checking a given condition.

1. What is the default value of the last parameter of VLOOKUP?

Answer:

TRUE/1 . It refers to finding the closest (approximate) match and assuming the table is sorted in ascending order. Whereas, FALSE/0 refers to exact match.

1. How to find duplicates in a column?

Answer:

Use CONDITIONAL FORMATTING to highlight duplicate values. OR use COUNTIF function as shown below. For example, values are stored in cells D4:D7.

*=COUNTIF(D4:D7,D4)*

Apply filter on the column wherein you applied the COUNTIF function and select values greater than 1.

1. How to insert a drop down?

Answer:

Go to Data tab >> Select Data Validation. Another way to insert a drop down is to enable Developer tab and Insert Combo box.

1. What's a pivot table and how's it used?

Answer:

"A pivot table can summarize large amounts of data. It allows you to analyze numerical data thoroughly and answer unanticipated questions, and present them in a user-friendly way. I once created a pivot table for a company that wanted to figure out which of their brands was producing the best sales over the course of two months. The pivot table had Column A showing the brands and Row 1 showing the months of October and November. I extracted information to compute total sales by brand per month. This allowed the company to identify which of its subsidiaries performed best."

1. Can you protect cells of a sheet from being copied?

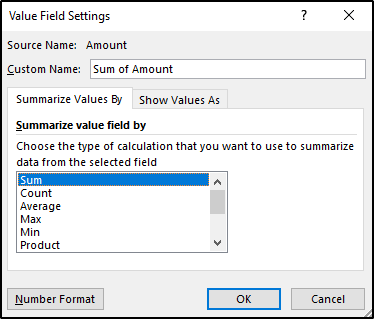
Answer:

Yes, you can do it by protecting the required cells or the complete sheet. In order to do this, follow the given steps:

1. Select the cells that you want to protect
2. Open up the Font window from the Home tab
3. From the Protection pane, select Protection and then check the Hidden box
4. Click on Review tab present in the Ribbon, and then select Protect sheet option (Excel will not hide the required cells unless you do this)
5. Specify a password (This will help you in unprotecting the sheet later)
6. How do you change the value field to show some other result other than the Sum?

Answer:

In order to change the value field to show results other than the Sum, right-click on the Sum of Amount values and then click on Value Field Settings. Here is the dialog box that you will see:



From here, you can select any value of your choice and then click on OK.

1. How to stop automatic sorting in PivotTables?

Answer:

Excel automatically sorts the data present in the Pivot Tables. In case you do not want Excel to do this, open the dropdown menu fro the Row Labels or the Column Labels, and then click on More Sort Options. You will see the Sort dialog box opening. Click on More Options and unselect the Sort automatically option.

1. What is Flash Fill?

Answer:

Flash Fill is an amazing tool that was added in Excel 2013 and is available in all version after that.

You can use Flash Fill in Excel to make data entry easy. It’s a smart tool that tries to identify patterns based on your data entry and does that for you.

Some simple examples of using Flash Fill could be to get the first name from the full name, get name initials, format phone numbers, etc.

11. Can you explain "data wrangling?"

Answer "Data wrangling involves cleaning raw data, structuring and enriching it into desired formats. This process effectively maps out large amounts of extracted data from several sources and turns it into a more usable format. It's useful because it improves the efficiency of all analysis that follows."

Q.12 While working in Excel, how can you clear all the formatting without removing cell content?"

Answer option found under the Home tab. Using this method will not remove any cell content. "

Q.13 What is Data Validation, screening, verification?

Answer Data validation, as the name suggests, is the process that involves determining the accuracy of data and the quality of the source as well. There are many processes in data validation but the main ones are data screening and data verification.

Data screening: Making use of a variety of models to ensure that the data is accurate and no redundancies are present.

Data verification: If there is a redundancy, it is evaluated based on multiple steps and then a call is taken to ensure the presence of the data item.

Q.14 Explain Data Cleaning in brief.

Answer Data Cleaning is also called Data Wrangling. As the name suggests, it is a structured way of finding erroneous content in data and safely removing them to ensure that the data is of the utmost quality. Here are some of the ways in data cleaning:

Removing a data block entirely

Finding ways to fill black data in, without causing redundancies

Replacing data with its mean or median values

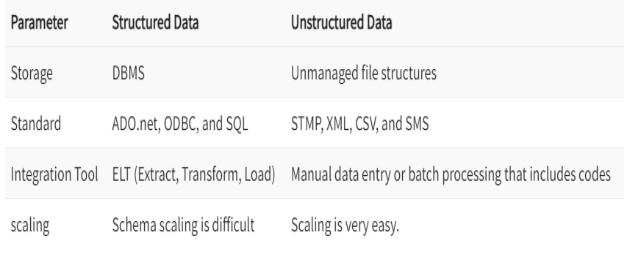
Making use of placeholders for empty spaces

Q.15 What is the use of a Pivot table?

Answer Pivot tables are one of the key features of Excel. They allow a user to view and summarize the entirety of large datasets simply. Most of the operations with Pivot tables involve drag-and-drop operations that aid in the quick creation of reports.

Q.16 Distinguish between structured and unstructured data

Answer



Q.17 What are four V’s of big data?

Answer Four V’s of big data are:

Velocity

Variety

Volume

Veracity

Q.18 List important fields or languages used by data engineer

Answer Here are a few fields or languages used by data engineer:

Probability as well as linear algebra

Machine learning

Trend analysis and regression

Hive QL and SQL databases

Q.19 What is the difference between Data Processing and Data Mining?

Answer Data Processing is the process of collecting and manipulating the items of a dataset to produce meaningful information.

Data Mining is a process of "mining" the data to uncover patterns. The term data mining is a misnomer (a wrong or inaccurate name or designation), because the goal is the extraction of patterns and knowledge from large amounts of data, not the extraction of data itself.

Q.20 Can you explain what a data analyst does?

"As a data analyst, I run, collect and process data for companies, which helps ensure accuracy of the information and helps managers make informed decisions. Another aspect of the job is implementing preventative measures, using critical thinking and attention to detail. "

Q.21 Which step of a Data Analysis project do you like the most?

Do know that it is completely normal to have a predilection toward certain tools and tasks over others. However, while performing data analysis, you will always be expected to deal with the entirety of the analytics life cycle, so make sure not to speak negatively about any of the tools or of the steps in the process of data analysis.

Finally, in this interview questions for the Data Analysts blog, we have to understand how to carefully approach this question and answer it to the best of our ability.

Q.22 Why is Python used for Data Cleaning in DS?

Data Scientists have to clean and transform the huge data sets in a form that they can work with. It is important to deal with the redundant data for better results by removing nonsensical outliers, malformed records, missing values, inconsistent formatting, etc.

Python libraries such as Matplotlib, Pandas, Numpy, Keras, and SciPy are extensively used for Data cleaning and analysis. These libraries are used to load and clean the data and do effective analysis. For example, a CSV file named “Student” has information about the students of an institute like their names, standard, address, phone number, grades, marks, etc.

Q.23 What are the popular libraries used in Data Science?

Below are the popular libraries used for data extraction, cleaning, visualization, and deploying DS models:

TensorFlow: Supports parallel computing with impeccable library management backed by Google.

SciPy: Mainly used for solving differential equations, multidimensional programming, data manipulation, and visualization through graphs and charts.

Pandas: Used to implement the ETL(Extracting, Transforming, and Loading the datasets) capabilities in business applications.

Matplotlib: Being free and open-source, it can be used as a replacement for MATLAB, which results in better performance and low memory consumption.

PyTorch: Best for projects which involve Machine Learning algorithms and Deep Neural Networks.

Q.24 Give the advantages of Cognos?

The advantages of Cognos are:

A) Planning

B) Analysis

C) Forecasting

D) Score carding

Q.25 List some of the skills and tools used by Business Analysts.

Answer this question by combining both the technical and non-technical tools/skills used by business analysts.

Technical skills/tool – MS Office Suite, Google Docs, database knowledge, ERP systems, SQL, and more.

Non-Technical/business Analysis skills – Documentation, requirement elicitation, business process management, and more.

Q.27 What are the tasks that are not part of a Business Analyst’s job?

Business Analyst is not part of the enlisted tasks:

Should not intend to organize the project team meetings.

Should not bother regarding the risks and issues tracker of a project.

Should not perform activities like testing (executing the TC’s), coding or programming.

Q.28 Explain the business analysis process flow.

Information gathering

Identify the key stakeholders

Identify the business objective

Determine the available options

Scope the definition

Define the delivery plan

Define the requirements of the project

Implementation and evaluation