بسم الله الرحمن الرحيم

Excel for Data Analysis – Full Course for Beginners

Project One

**Salaries of Data Science Jobs**

**(Jobs Calculator / Dashboard)**

**Content:**

* Overview.
* Part 1: Prepare Data for Analysis.
* Part 2: Analyze the Data.

**Written by:**

**Abobaker Ahmed Khidir Hassan**

Overview

About the Project

This course is an Excel course for Data Analysis. It’s starting from installing Excel up to creating a full project using DAX.

This project is the first full project in the course and it use the source data contains 16 columns and 32,762 rows about the Data Science Jobs Posts to create a calculator or mini dashboard that helps taking decision which job title must be focused in.

Lessons Have Been Taken:

* What is Excel?
* Formulas Intro
* Function Intro
* Logical Functions
* Math Functions
* Statistical Functions
* Array Formulas
* Lookup Function
* Text Functions
* Date and Time Functions
* Charts Intro
* Charts Advanced
* Charts Statistics
* Sparklines
* Tables
* Formatting
* Collaboration
* Project #1: Build Dashboard
* Project #1: Share Projects

Data

The original data is consists of 16 columns and 32,762 rows/records.

| **Column Name** | **Description** | **Inferred Data Type** |
| --- | --- | --- |
| job\_title\_short | Shortened version of job title | string |
| job\_title | Full job title and department | string |
| job\_location | City and state of the job | string |
| job\_via | Platform where the job was posted | string |
| job\_schedule\_type | Full-time, part-time, etc. | string or categorical |
| job\_work\_from\_home | Whether the job allows remote work | boolean |
| search\_location | Location used for searching jobs (possibly user input) | string |
| job\_posted\_date | Date since the job was posted (appears to be in float days format) | float or int |
| job\_no\_degree\_mention | Whether the job mentions that a degree is *not* required | boolean |
| job\_health\_insurance | Whether health insurance is offered | boolean |
| job\_country | Country where the job is located | string |
| salary\_rate | Salary time unit (e.g., year, hour) | string (categorical) |
| salary\_year\_avg | Average annual salary | float |
| salary\_hour\_avg | Average hourly salary (may be null) | float or null |
| company\_name | Employer offering the job | string |
| job\_skills | List of required skills | list[string] or string |

Preparing Data for Analysis

In this section, I will clean and applying many concepts about the data preprocessing as what we will se.

Step 1: Create a Copy of the Data

This is my first step before doing anything, and we know how it’s important to save the original data in a safe place.

I named the new worksheet WorkSpace.

Step 2: Formatting Data as a Table

To make referencing more efficient and reduce the referencing error those maybe happen in the cell referencing.

I named it jobs.

Step 3: Formatting

I set the format/data types of the fields then removing spalling errors, blanks to make sure the analysis will be applied in a clean data.

Note:

I note that there is some uncompleted data in some columns such as:

job\_location, job\_via, job\_schedule\_type, and job\_skills.

And I note 10636 of the salaries are yearly salaries, and the others (22036) are hourly salaries.

All of these will be in mind during the analysis.

Analyze the Data

In this section, I will start getting insights from that data.

Step 1: What will help someone choose?

Let as thinking for a moment, if you want to chose a job title to focus on, what you will depend on in your chose?

Maybe your passion, but I don’t know what it is.

Maybe the average salaries? Yes I think that is a good info.

But salaries are different based on your country? I think the country will be a good filter too.

Are you busy? If yes, I think you need a part-time job, is it? So I will add the schedule type as another filter.

What about the skills? Yes it is a good info you need to know.

So.. If you choose and ready, and you need to find a job, so I will let the top platform and the best time for searching on the jobs posts about your dream job.

s is my first step before doing anything, and we know how it’s important to save the original data in a safe place.

I named the new worksheet WorkSpace in a new workbook.

Step 2: Creating the Validation and the Calculator Worksheets

To make sure the info those gotten from the raw data are valid, I will apply some conditions about what data will be included in the analysis, I will perform that in the Validation sheet.

The Calculator sheet will hold the final dashboard, I will create it from now to build it section by section.

Step 3: Info Should Be Known

As we say, two of the most info you should know before choosing the job, is the salary and the job counts, so I putted them in the calculation sheet as KPIs cards.

For salaries, I’ll use the median to reduce the effectiveness of the outliers.

Step 4: Titles

Let’s creating the salary sheet to analyze the salaries of each job title

Before that I validate the job title using the unique and sore functions in the validation sheet. I sort them depend on the job counts.

It that sheet I bring the titles from the validation sheet, then use the average of the yearly salary after filtering the blanks (If not, they will be calculated as zeros!), and filter it depend on the job title too.

I created a dropdown menu (filter) in the calculator sheet that makes available to select a single job title and update calculating the job counts and median salary depended on it.

Step 5: Countries

In this step, I created a sheet named country and import the countries from the validation sheet, and that coming after importing it from the work space with the unique and sort function after filtering missing values.

I also calculate the median salary of each country depending on the job title after filtering the blanks.

I created a dropdown menu in the calculator sheet that makes available to select a single country and make the calculation of the count of jobs and the median salary depended on both, the job title and the country.

Step 6: Types

Now, I created a sheet named type and import the types from the validation sheet, and that coming after importing them from the work space with the unique function after filtering them into the main four types.

I created a dropdown menu in the calculator sheet that makes available to select a single country and update the calculation of the count of jobs and the median salary to be depended on the job title, the country, and the job type.

Step 7: Top Platforms

After finishing all of these calculations, I created a new sheet named platform where I will find the top platform has posts about the selected job title, country, and job type.

I imported the unique platforms names in the validation, then I uploaded them in the platform sheet.

I calculated the top one depending on the title, country, and type, then I displayed the top one in the calculator sheet.

Step 8: Time

To find a good job, you have to know when is the best time to search, and how to decided to agree a median offer or reject it to find another one.

To do that I created a sheet named time, and putted the counts of jobs of each month depend on the previous factors.

I also add the table shows the count of posts in the hours of a day to know what is the best time to stay online looking for a job.