

WUZZUF Job Postings Analytics

Agenda

- Data Cleaning
- Data Analysis and visualization

Data Understanding and Cleaning

The first thing I did when I downloaded the data that I started to explore what is the data contain, what is the domain of the data I'm dealing with, and trying to know and discover many other things that definitely would help any data analyst want to make the data spill it's secrets.

In this phase I used Power Query Editor to help me explore, clean, organize, and handle that data.

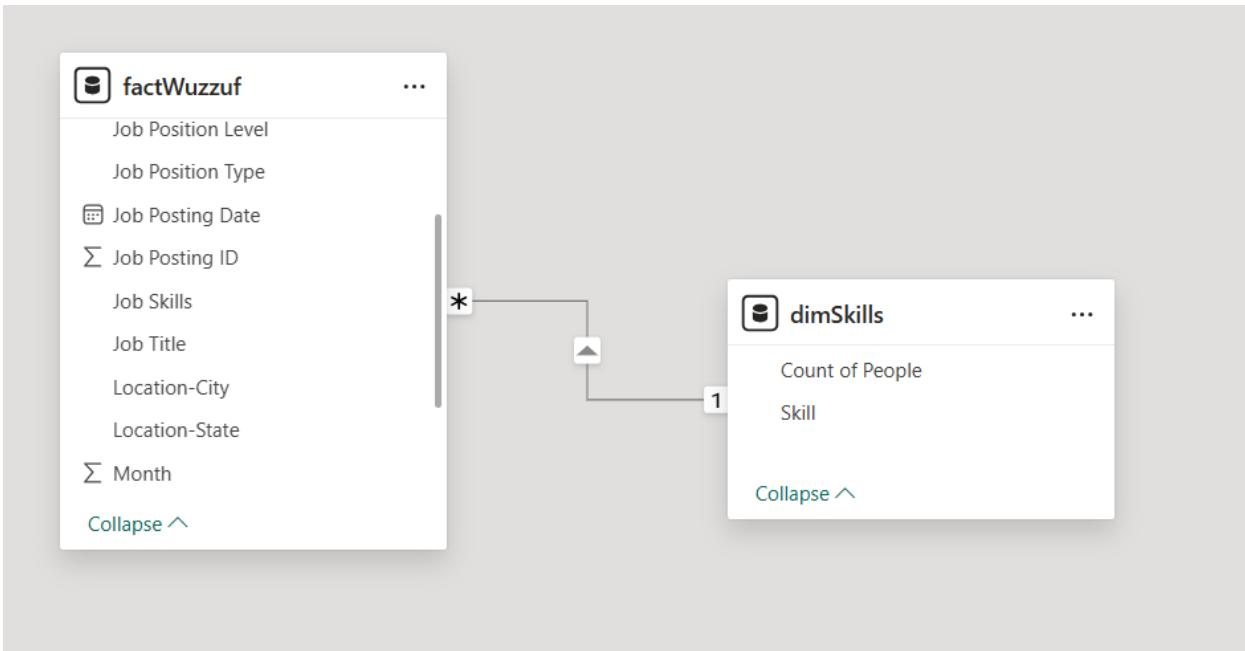
And here's a list of the issues in the data and how I handled them:

- I noticed a useless index column so I deleted it and checked the data types
- I have seen so many inconsistencies, errors, and blanks (70 % +) of the entire column of "minimum pay", "maximum pay", "pay rate", "number of applicants", "additional info about job", and "Range of Salary" so I removed them.
- Then I splitted the location column into two columns after eyeballing and noticing that it contains information about "city,state" and also there were some inconsistencies with the column where some rows contain California and others (which are the most of the table) typed as an abbreviation so I changed some whole names of states into the abbreviation to maintain consistency by using the replace values feature. Also regarding the "location" column I converted the values that contain United or null into United States for clarity and better map visualization
- After that, I added two new columns which are Year, and Month from the date column.
- Last but not least I duplicated the entire table to split, unpivot, and aggregate the skills by the number of jobs.

Source	⚙
Promoted Headers	⚙
Changed Type	⚙
Removed Columns	⚙
Replaced Value	⚙
Merged Columns	⚙
Split Column by Delimiter	⚙
Changed Type1	⚙
Filtered Rows	⚙
Renamed Columns	⚙
Filtered Rows1	⚙
Replaced Value1	⚙
Replaced Value2	⚙
Merged Columns1	⚙
Replaced Value3	⚙
Replaced Value4	⚙
Removed Columns1	⚙
Filtered Rows2	⚙
Filtered Rows3	⚙
Inserted Year	⚙
✖ Inserted Month	⚙

	A _C Job Skills.2.1.2	A _C Job Skills.2.1.4	A _C Job Skills.2.1.6	A _C Job Skills.2.1.8	A _C Job Skills.2.1.10	A _C Job Skills.2.1.12	A _C Job Skills.2.1.14
1	agile	excel		null	null	null	null
2	database	pl/sql	cloud	javascript	agile	node.js	git
3	data_lake	database	python	hadoop	ibm	oracle	microsoft
4	python	programming	etl	postgresql	linux	sql	nosql
5	r	scala	powershell	c++	java	python	tableau
6	cloud	programming	microsoft	aws	go	azure	
7	ci/cd	cloud	python	programming	linux	docker	
8	cloud	power.bi	agile	bi	tableau	etl	sql
9	javascript		null	null	null	null	null
10	backend	java		null	null	null	null
11	scrum	java	aws	agile		null	null
12	r	python	programming	sql	ai	machine_learning	
13	javascript	agile	python	scrum	programming	php	c

	A ^B _C Attribute	A ^B _C Skill
1	Job Skills.2.1.22	agile
2	Job Skills.2.1.10	agile
3	Job Skills.2.1.26	agile
4	Job Skills.2.1.4	agile
5	Job Skills.2.1.20	agile
6	Job Skills.2.1.18	agile
7	Job Skills.2.1.6	agile
8	Job Skills.2.1.8	agile
9	Job Skills.2.1.24	agile
10	Job Skills.2.1.30	agile
11	Job Skills.2.1.2	agile
12	Job Skills.2.1.16	agile
13	Job Skills.2.1.32	agile
14	Job Skills.2.1.12	agile
15	Job Skills.2.1.14	agile
16	Job Skills.2.1.28	agile
17	Job Skills.2.1.10	agile/scrum
18	Job Skills.2.1.14	agile/scrum
19	Job Skills.2.1.6	agile/scrum
20	Job Skills.2.1.4	agile/scrum
21	Job Skills.2.1.12	agile/scrum
22	Job Skills.2.1.2	agile/scrum
23	Job Skills.2.1.16	agile/scrum
24	Job Skills.2.1.8	agile/scrum
25	Job Skills.2.1.38	ai
26	Job Skills.2.1.26	ai
27	Job Skills.2.1.32	ai
28	Job Skills.2.1.36	ai



Data Analysis and Visualization

After the phase of exploring and cleaning the data. It's time to analyze by writing dax to create the appropriate measures that would help me in my analysis.

And visualize the insights that would help the business understand more about the trends of jobs, number of postings, the average years of experience required for most jobs, which companies hire most?, what is the trend of job postings over time?, and many more questions that are displayed in a nice dashboard that contain all these insights and knowledge.

```
Total Postings = COUNT(factWuzzuf[Job Posting ID])
```

```
Average Years of Experience = average(factWuzzuf[Years of Experience])
```

```
Most Job Offering Company = CALCULATE [MAXX (VALUES(factWuzzuf[Company Name]),CALCULATE (COUNT(factWuzzuf[Job Posting ID])))])]
```

