THE FUTURE OF WORK : DATA ANALYSIS OF GLASSDOOR JOBS

Category: Data Analytics

From Aditya College Of Engineering.

Team ID : LTVIP2023TMID04038

Team Leader : Karri Yamuna Ramalakshmi Ramya

Team Member :Katta Chinna Venkata Balaji

Team Member : Kintali Leela Prasad

Team Member : Koka Naga Pavan

Team Member : Tirumalasetti Koti Surya Durga Ramakumar

INDEX :

1. INTRODUCTION:

Abstract:

The "Future of work : Data Analysis of Glassdoor Jobs" project aim is to conduct an analysis of Glassdoor job postings to gain insights into current and emerging job market trends, identify in-demand skills and experience, and understand how employers can improve their employer branding and reputation to attract and retain top talent.

1.1 Project Description:

Job analysis is a systematic procedure to analyze the requirements for the job role and job profile. Glassdoor is a website and online platform that provides information about jobs, salaries, and companies. Job analysis is a systematic approach to defining the job role, description, requirements, responsibilities, evaluation, etc. It helps in finding out the required level of education, skills, knowledge, training, etc for the job position. It also depicts the job worth i.e. measurable effectiveness of the job and contribution of job to the organization. Thus, it effectively contributes to setting up the compensation package for the job position.

Lack of analysis of Glassdoor jobs can result in limited understanding of job market trends, difficulty in finding relevant job opportunities, inability to attract and retain top talent, and lack of insight into company branding and reputation.

The purpose of this project is to conduct an analysis of Glassdoor job postings to gain insights into current and emerging job market trends, identify in-demand skills and experience, and understand how employers can improve their employer branding and reputation to attract and retain top talent.

1.2 Purpose of the project:

The project "Future of work:Data Analysis of Glassdoor Jobs" through data analytics can have various purpose such that would likely involve analyzing data from Glassdoor job postings to gain insights into trends related to the future of work.The project could serve various purposes, such as:

1. **Identifying emerging job roles and skills:** Analyzing job postings can help identify new job titles and skills that are becoming more relevant in the job market, providing valuable information for job seekers and employers.
2. **Predicting industry changes:** By analyzing data from various job postings, patterns and trends may emerge, allowing analysts to make predictions about shifts within specific industries or sectors.
3. **Understanding job market demands:** The project might aim to understand what employers are looking for in candidates, helping job seekers tailor their skills and profiles accordingly.

4. **Uncovering remote work opportunities:** With the rise of remote work, analyzing job postings can reveal remote-friendly job opportunities and help understand how remote work is impacting the job market.
5. **Exploring diversity and inclusion aspects:** The analysis could shed light on diversity and inclusion efforts in the workplace, such as the representation of various demographics in different job roles.
6. **Supporting policy decisions:** The insights gained from the project might be used by policymakers and organizations to make informed decisions about workforce development and labor market strategies.

2. LITERATURE SURVEY

A literacy survey for Data Analysis of Glassdoor Jobs involves reviewing multiple job roles in a particular domain offered by a particular organization belonging to a given industry and sector. Job analysis defines the organization of jobs within a job family. It allows units to identify paths of job progression for employees interested in improving their opportunities for career advancement and increasing compensation.

Existing Problems:

There are several existing problems and challenges can arise in the future of work data analysis of the Glassdoor jobs project. It's important to be aware of these issues to ensure the accuracy, fairness, and reliability of the analysis. Some of the key problems include:

1. **Data Quality and Availability:** The quality of data available on Glassdoor may vary, with some job listings lacking detailed information or having outdated content. Incomplete or inaccurate data can lead to biased or misleading analyses.
2. **Selection Bias:** The data available on Glassdoor may not represent the entire job market, as it mainly reflects positions advertised by certain companies and industries. This selection bias can affect the generalizability of the findings.
3. **Privacy and Ethical Concerns:** Glassdoor data contains personal information about employees and company reviews, raising privacy and ethical concerns. Ensuring data anonymity and complying with data protection regulations is crucial.
4. **Text Data Complexity:** Analyzing unstructured text data from job descriptions and reviews can be challenging due to its complexity and ambiguity. NLP techniques may struggle to capture the full context and nuance in the text.
5. **Sentiment Analysis Accuracy:** Sentiment analysis may not always accurately interpret the sentiment expressed in reviews, leading to potential misinterpretations of employee satisfaction.
6. **Skills and Job Title Standardization:** Different companies may use various terms to describe similar skills or job titles, making it difficult to compare and group data consistently.
7. **Temporal Dynamics:** The job market is continuously evolving, and trends may change rapidly. The analysis should consider the temporal dynamics to provide up-to-date and relevant insights.
8. **Overfitting in Predictive Models:** While predictive analytics can offer valuable insights, overfitting the models to historical data might lead to inaccurate predictions in dynamic job markets.
9. **Lack of Context:** The data available on Glassdoor may lack contextual information about market conditions, industry trends, or external factors that can influence job trends.

10. **Interpretability of AI Models:** When using AI and machine learning models, the interpretability of results becomes crucial, especially for decision-making purposes. Black-box models may be challenging to explain and understand.

2.2 PROPOSED SOLUTION:

The proposed solution for "The Future of work:data analysis of the Glassdoor jobs" project are:

1. **Data Collection:** Gather job-related data from Glassdoor's API or web scraping. This data could include job titles, salaries, location, company ratings, job descriptions, and other relevant information.
2. **Data Cleaning:** Process the collected data to remove duplicates, handle missing values, and standardize formats. This step ensures that the data is in a usable form for analysis.
3. **Exploratory Data Analysis (EDA):** Perform exploratory analysis to understand the dataset's characteristics, identify trends, and discover insights. Visualization techniques can help in presenting the findings effectively.
4. **Sentiment Analysis:** Use Natural Language Processing (NLP) techniques to analyze job reviews and comments to gauge employee satisfaction, sentiments, and potential areas for improvement in companies.
5. **Salary Analysis:** Compare salaries across different job titles, industries, and locations. Identify the factors influencing salary variations and provide recommendations based on the findings.
6. **Company Ratings and Reviews**: Analyze company ratings and reviews to identify top-rated companies and factors contributing to positive/negative reviews.
7. **Job Market Trends:** Analyze job postings over time to understand the evolving job market, popular skills, and emerging job opportunities.
8. **Predictive Analytics:** Utilize machine learning models to predict future job trends, salary changes, or company ratings based on historical data.
9. **Data Visualization:** Create interactive and informative data visualizations to communicate the findings effectively to stakeholders.
10. **Recommendations:** Based on the analysis, provide actionable recommendations for job seekers, companies, and policymakers to enhance the future of work.

2.3. SOCIAL AND BUSINESS IMPACT :

The project of "Future of Work : Data Analysis of Glassdoor Jobs" can have significant social and business impacts that are

2.3.1 Social Impact:

This project can help job seekers make more informed decisions about their careers and negotiate for better compensation and working conditions. This can ultimately contribute to greater economic mobility and reduce income inequality.

1. **Skill demand:** Data analysis skills are becoming increasingly essential in various industries, leading to a shift in the required skill set for the workforce. Employees with data analysis expertise may have better job prospects and opportunities for career growth.
2. **Job displacement and creation:** Automation and data-driven technologies could lead to the displacement of certain jobs that can be automated. However, this also creates new job opportunities in data analysis, data science, and related fields.
3. **Workforce diversity and inclusion:** Data analysis can help identify and address biases in hiring practices, promoting diversity and inclusion in the workplace by ensuring fair recruitment and evaluation processes.
4. **Work-life balance:** With improved data analysis, companies can gain insights into employee performance, leading to more effective strategies for promoting work-life balance and well-being.

2.3.2 Business Impact:

It can help to improve retention rates, reduce turnover costs, and increase productivity. An analysis of Glassdoor jobs can provide insights into what employees value most, helping employers to create a better work environment that attracts and retains top talent.

1. **Data-driven decision making:** Organizations can leverage data analysis to make informed and strategic decisions, leading to improved efficiency, cost reduction, and competitive advantage.
2. **Personalization and customer satisfaction:** By analyzing customer data, businesses can offer personalized products and services, leading to higher customer satisfaction and loyalty.
3. **Operational efficiency:** Data analysis can optimize processes, supply chains, and resource allocation, resulting in improved productivity and reduced waste.
4. **Talent acquisition and retention:** Data analysis helps in identifying top talent, matching them with suitable roles, and implementing strategies to retain valuable employees.
5. **Market trends and competition:** By analyzing market trends and competitors' activities, businesses can adapt their strategies to stay competitive and identify emerging opportunities.

Remember that the future of work is influenced by various factors, and data analysis is just one of them. The ethical and responsible use of data is crucial to ensure positive outcomes and minimize potential negative impacts on individuals and society as a whole.

3.THEORETICAL ANALYSIS:

3.1 Block diagram:

FIG:01

3.2 Dataset:

The dataset must be available on github here and the data contains all the meta information regarding the columns described in the CSV files. The name of file is Cleaned_DS_Jobs.csv.

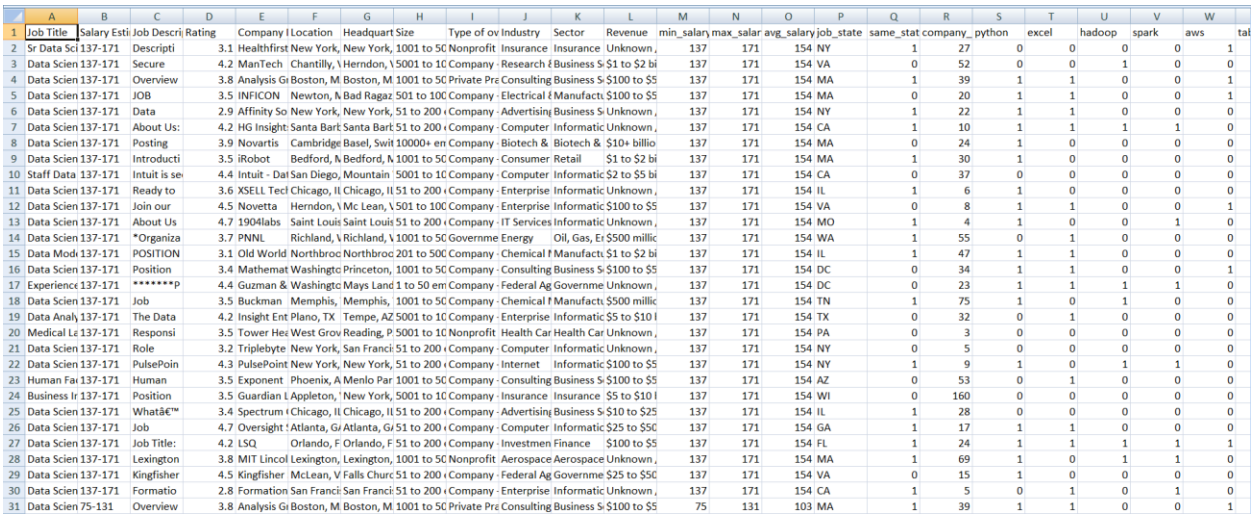
Description for Cleaned_DS_Jobs.csv:

The file Cleaned_DS_Jobs.csv contains 660 rows. Each row corresponds to a record of state with details and marks in respective subjects.

The columns are:

Job Title: Title of the job posting.
Salary Estimation: Salary range for that particular job.
Job Description: This contains the full description of that job,
Rating: Rating of that post.
Company: Name of company.
Location: Location of the company.
Headquarter: Location of the headquarter.
Size: Total employee in that company.
Type of ownership: Describes the company type i.e non-profit/public/private farm etc.
Industry, Sector: Field applicant will work in.
Revenue: Total revenue of the company.

min_salary,max_salary,avg_salary: Refers to the minimum, maximum and average salary for that post.
job_state: State where the applicant will work.
same_state: Same state as headquarter or not(Boolean).
company_age: Age of company.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Job Title	Salary Esti	Job Descri	Rating	Company I	Location	Headquart	Size	Type of ov	Industry	Sector	Revenue	min_salary	max_salar	avg_salary	job_state	same_stat	company_	python	excel	hadoop	spark	aws
2	Sr Data Sci	137-171	Descripti	3.1	Healthfirst	New York,	New York,	1001 to 50	Nonprofit	Insurance	Insurance	Unknown /	137	171	154	NY	1	27	0	0	0	0	1
3	Data Scien	137-171	Secure	4.2	ManTech	Chantilly, \	Herndon, \	5001 to 10	Company -	Research &	Business S	$1 to $2 bi	137	171	154	VA	0	52	0	0	1	0	0
4	Data Scien	137-171	Overview	3.8	Analysis Gi	Boston, M	Boston, M	1001 to 50	Private Pra	Consulting	Business S	$100 to $5	137	171	154	MA	1	39	1	1	0	0	1
5	Data Scien	137-171	JOB	3.5	INFICON	Newton, N	Bad Ragaz	501 to 100	Company -	Electrical &	Manufactu	$100 to $5	137	171	154	MA	0	20	1	1	0	0	1
6	Data Scien	137-171	Data	2.9	Affinity So	New York,	New York,	51 to 200	Company -	Advertisin	Business S	Unknown /	137	171	154	NY	1	22	1	1	0	0	0
7	Data Scien	137-171	About Us:	4.2	HG Insight:	Santa Barb	Santa Barb	51 to 200	Company -	Computer	Informatic	Unknown /	137	171	154	CA	1	10	1	1	1	1	0
8	Data Scien	137-171	Posting	3.9	Novartis	Cambridge	Basel, Swit	10000+ em	Company -	Biotech &	Biotech &	$10+ billio	137	171	154	MA	0	24	1	0	0	0	0
9	Data Scien	137-171	Introducti	3.5	iRobot	Bedford, N	Bedford, N	1001 to 50	Company -	Consumer	Retail	$1 to $2 bi	137	171	154	MA	1	30	1	0	0	0	0
10	Staff Data	137-171	Intuit is se	4.4	Intuit - Dat	San Diego,	Mountain	5001 to 10	Company -	Computer	Informatic	$2 to $5 bi	137	171	154	CA	0	37	0	0	0	0	0
11	Data Scien	137-171	Ready to	3.6	XSELL Tecl	Chicago, IL	Chicago, IL	51 to 200	Company -	Enterprise	Informatic	Unknown /	137	171	154	IL	1	6	1	0	0	0	0
12	Data Scien	137-171	Join our	4.5	Novetta	Herndon, \	Mc Lean, \	501 to 100	Company -	Enterprise	Informatic	$100 to $5	137	171	154	VA	0	8	1	1	0	0	1
13	Data Scien	137-171	About Us	4.7	1904labs	Saint Louis	Saint Louis	51 to 200	Company -	IT Services	Informatic	Unknown /	137	171	154	MO	1	4	1	0	0	1	0
14	Data Scien	137-171	*Organiza	3.7	PNNL	Richland, \	Richland, \	1001 to 50	Governme	Energy	Oil, Gas, Er	$500 millic	137	171	154	WA	1	55	0	1	0	0	0
15	Data Mod	137-171	POSITION	3.1	Old World	Northbroo	Northbroo	201 to 500	Company -	Chemical I	Manufactu	$1 to $2 bi	137	171	154	IL	1	47	1	1	0	0	0
16	Data Scien	137-171	Position	3.4	Mathemat	Washingto	Princeton,	1001 to 50	Company -	Consulting	Business S	$100 to $5	137	171	154	DC	0	34	1	1	0	0	1
17	Experience	137-171	********P	4.4	Guzman &	Washingto	Mays Land	1 to 50 em	Company -	Federal Ag	Governme	Unknown /	137	171	154	DC	0	23	1	1	1	1	0
18	Data Scien	137-171	Job	3.5	Buckman	Memphis,	Memphis,	1001 to 50	Company -	Chemical I	Manufactu	$500 millic	137	171	154	TN	1	75	1	0	1	0	0
19	Data Analy	137-171	The Data	4.2	Insight Ent	Plano, TX	Tempe, AZ	5001 to 10	Company -	Enterprise	Informatic	$5 to $10 I	137	171	154	TX	0	32	0	1	0	0	0
20	Medical La	137-171	Responsi	3.5	Tower Hea	West Grov	Reading, P	5001 to 10	Nonprofit	Health Car	Health Car	Unknown /	137	171	154	PA	0	3	0	0	0	0	0
21	Data Scien	137-171	Role	3.2	Triplebyte	New York,	San Franci	51 to 200	Company -	Computer	Informatic	Unknown /	137	171	154	NY	0	5	0	0	0	0	0
22	Data Scien	137-171	PulsePoin	4.3	PulsePoint	New York,	New York,	51 to 200	Company -	Internet	Informatic	$100 to $5	137	171	154	NY	1	9	1	0	1	1	0
23	Human Fa	137-171	Human	3.5	Exponent	Phoenix, A	Menlo Par	1001 to 50	Company -	Consulting	Business S	$100 to $5	137	171	154	AZ	0	53	0	1	0	0	0
24	Business Ir	137-171	Position	3.5	Guardian L	Appleton,	New York,	5001 to 10	Company -	Insurance	Insurance	$5 to $10 I	137	171	154	WI	0	160	0	0	0	0	0
25	Data Scien	137-171	Whatâ€™	3.4	Spectrum (	Chicago, IL	Chicago, IL	51 to 200	Company -	Advertisin	Business S	$10 to $25	137	171	154	IL	1	28	0	0	0	0	0
26	Data Scien	137-171	Job	4.7	Oversight !	Atlanta, GA	Atlanta, GA	51 to 200	Company -	Computer	Informatic	$25 to $50	137	171	154	GA	1	17	1	1	0	0	0
27	Data Scien	137-171	Job Title:	4.2	LSQ	Orlando, F	Orlando, F	51 to 200	Company -	Investmen	Finance	$100 to $5	137	171	154	FL	1	24	1	1	1	1	1
28	Data Scien	137-171	Lexington	3.8	MIT Lincol	Lexington,	Lexington,	1001 to 50	Nonprofit	Aerospace	Aerospace	Unknown /	137	171	154	MA	1	69	1	0	1	1	0
29	Data Scien	137-171	Kingfisher	4.5	Kingfisher	McLean, V	Falls Churc	51 to 200	Company -	Federal Ag	Governme	$25 to $50	137	171	154	VA	0	15	1	0	0	0	0
30	Data Scien	137-171	Formatio	2.8	Formation	San Franci	San Franci	51 to 200	Company -	Enterprise	Informatic	Unknown /	137	171	154	CA	1	5	0	1	0	1	0
31	Data Scien	75-131	Overview	3.8	Analysis Gi	Boston, M	Boston, M	1001 to 50	Private Pra	Consulting	Business S	$100 to $5	75	131	103	MA	1	39	1	1	0	0	1

FIG:02

3.3 Connect DB2 with Cognos:

To connect IBM DB2 and IBM Cognos Analytics, you'll need to set up a data source connection in Cognos Analytics to access the DB2 database.

Here's a step-by-step guide on how to do it:

Ensure Prerequisites:

- Make sure you have the necessary credentials (username and password) to access the DB2 database.Obtain the DB2 database connection details, including the hostname or IP address, port number, and database name
- Launch IBM Cognos Analytics
- Log in to IBM Cognos Analytics with your credentials.

Access the Administration Console:

- In the Cognos Analytics user interface, click on the "Hamburger" menu icon (three horizontal lines) in the top-left corner.
- From the menu, select "Administration."
- In the Administration Console, expand the "Configuration" section in the left pane.
- Click on "Data Sources" under "Configuration."

Add a New Data Server:

- On the "Data Server" page, click the "Add" button to create a new data source connection.
- In the "Select the type of data source" window, choose "IBM DB2" from the list of available data sources.
- Click "Next" to proceed.

Provide Connection Details:

- Fill in the required connection details for the DB2 database:
- Enter a name for the data source connection (e.g., "My DB2 Connection").
- Specify the hostname or IP address of the DB2 server.
- Enter the port number on which DB2 is listening.
- Provide the database name.
- Input your DB2 username and password for authentication.

Test the Connection:

- Click the "Test" button to verify if the connection to the DB2 database is successful. Cognos Analytics will attempt to establish a connection using the provided details.
- If the test is successful, click "OK" to save the date of the Data Server Source connection.

Publish the Data Server:

- After saving the data source, click on the "More Actions" button (three dots) next to the data source you created.
- Select "Publish" to make the data source available for use in reports and dashboards.

Create Reports and Dashboards:

- With the data source connection established, you can now use IBM Cognos Analytics to create reports and dashboards by selecting the DB2 data source as a data provider.
- That's it! You have now connected IBM DB2 and IBM Cognos Analytics, allowing you to leverage the data from the DB2 database to generate meaningful insights and visualizations within the Cognos Analytics platform.

3.4 Prepare The Data For Visualization :

To prepare the data for visualization in IBM Cognos Analytics, you need to perform certain data preparation steps to ensure the data is in the right format and structure for effective visualization. Here's a guide to prepare the data:

Data Source Connection:

- Connect IBM Cognos Analytics to the data source where your data is stored. This could be a relational database like IBM DB2, a data warehouse, Excel files, or other data repositories.

Data Import:

- Import the required data into Cognos Analytics. This involves creating a new data module or importing data directly from the data source. The data module allows you to combine data from different sources if needed.

Data Quality Check:

- Perform data quality checks to identify and handle any missing values, anomalies, or inconsistencies in the data. Clean the data by handling missing values appropriately (e.g., inputting, removing, or leaving).

Data Transformation:

- Transform the data as required for visualization. This may involve aggregating data, calculating new measures or metrics, creating calculated fields, and applying data formatting (e.g., date formatting, number formatting).

Creating Data Groups and Hierarchies:

- Create data groups and hierarchies to organize and structure the data for easier visualization. This is particularly useful for organizing data in dimensions like time (year, quarter, month) or geographic regions (country, city).

Sorting Data:

- Sort the data appropriately to present it in a meaningful and logical order in the visualizations.

Data Aggregation:

- If necessary, aggregate data to higher levels for summary and aggregation visualizations like charts and graphs.

Data Preview and Validation:

- Preview the data to ensure that it is prepared correctly and that all the required transformations and calculations have been applied accurately.

Save and Organize:

- Save the prepared data module in IBM Cognos Analytics, ensuring it is properly organized in the appropriate folders for easy access and reuse.

Once the data is prepared in IBM Cognos Analytics, you can create a variety of visualizations such as charts, graphs, tables, and maps to gain insights from the data. Data preparation is a crucial step to ensure that the visualizations accurately represent the underlying data and help users make informed decisions based on the insights gained.

3.5 Hardware / Software designing:

HARDWARE USED:
LAPTOP, MOBILE PHONE

SOFTWARE USED:
IBM COGNOS, ANACONDA, PYTHON,FLASK INTEGRATION

4. RESULT :

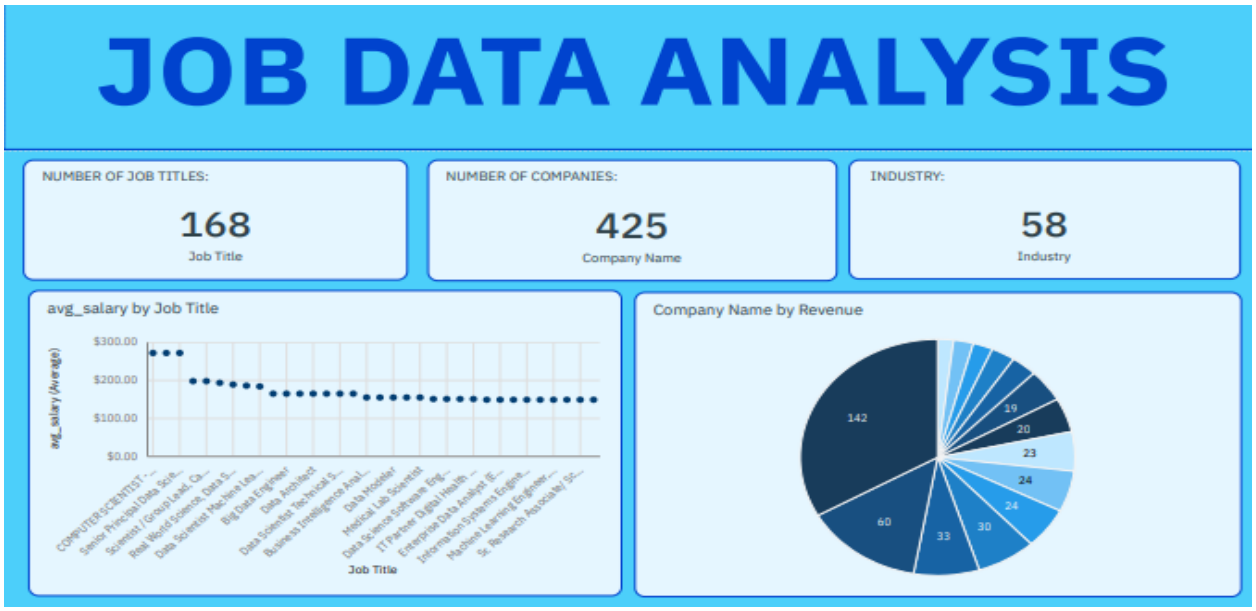


FIG:03

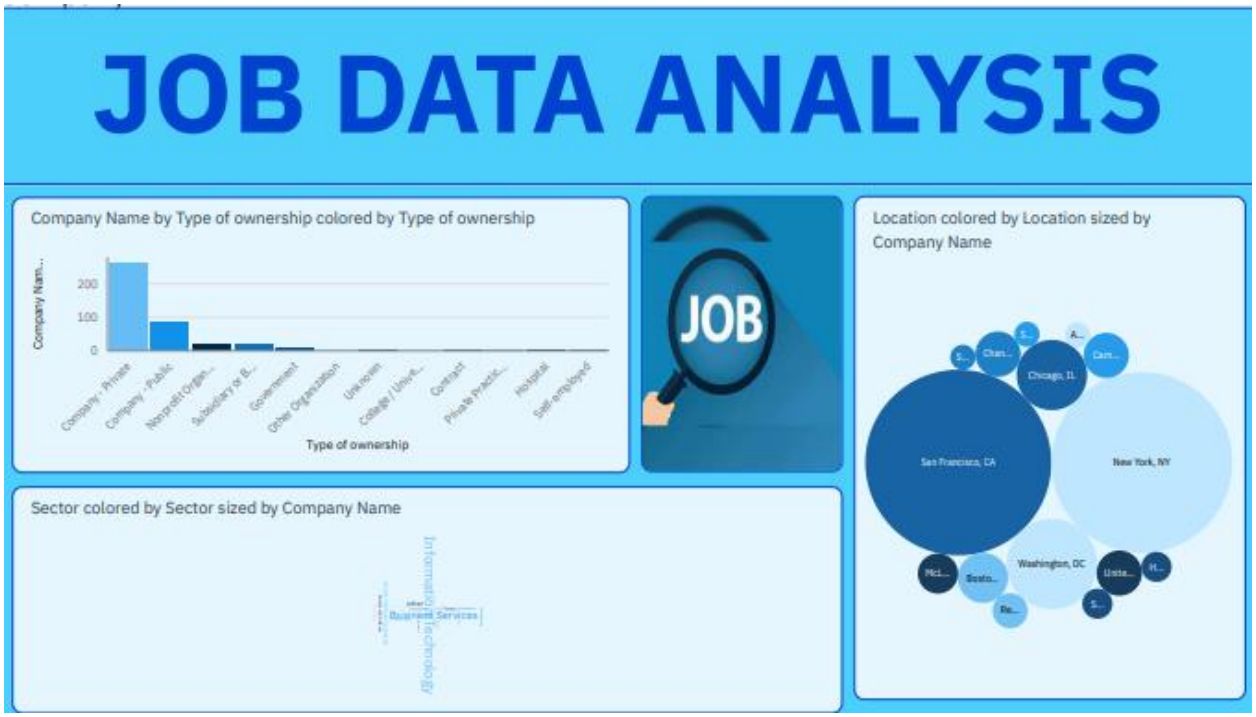


FIG:04

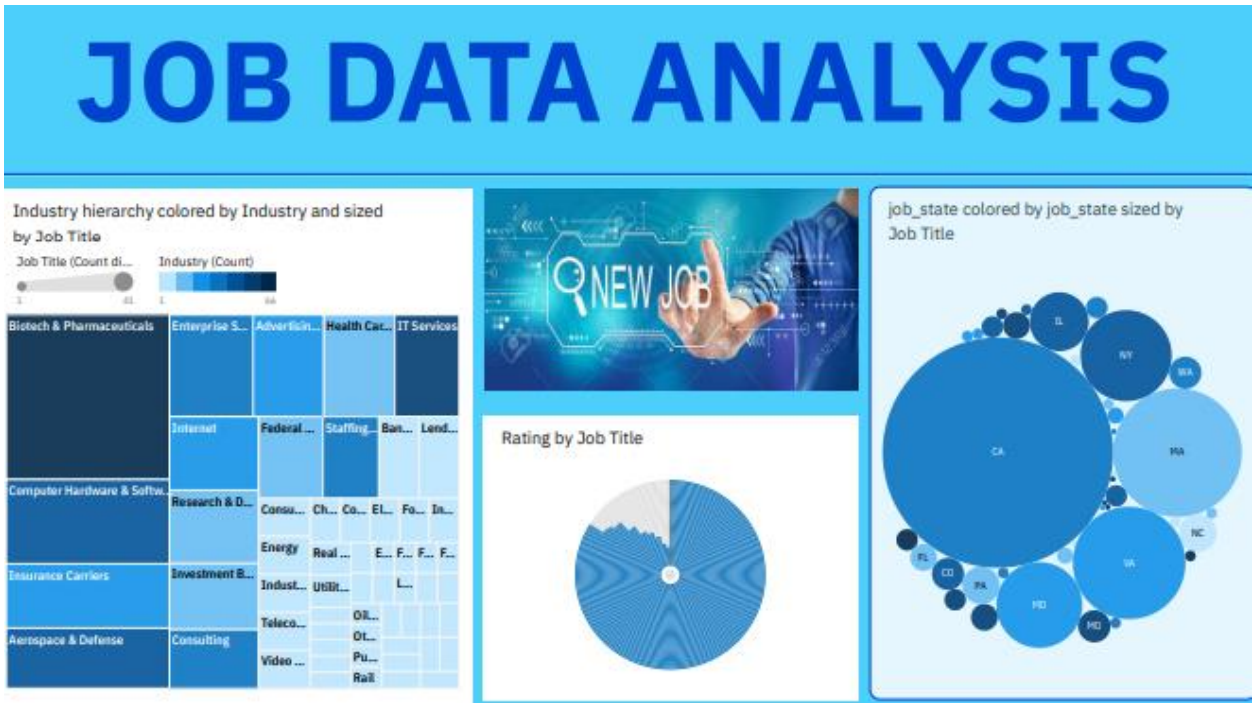

JOB DATA ANALYSIS
Industry hierarchy colored by Industry and sized by Job Title
Job Title (Count di...
Industry (Count)
Biotech & Pharmaceuticals
Enterprise S...
Advertisin...
Health Car...
IT Services
Internet
Federal ...
Staffing ...
Ban...
Lend...
Computer Hardware & Softw...
Research & D...
Consu...
Ch...
Co...
El...
Fo...
In...
Energy
Real ...
E...
F...
F...
F...
Insurance Carriers
Investment B...
Indust...
Utilit...
L...
Teleco...
Oil...
Ot...
Aerospace & Defense
Consulting
Video ...
Pu...
Rail
NEW JOB
Rating by Job Title
job_state colored by job_state sized by Job Title
IL
NY
MA
CA
NC
VA
FL
CO
PA
MD

FIG:05

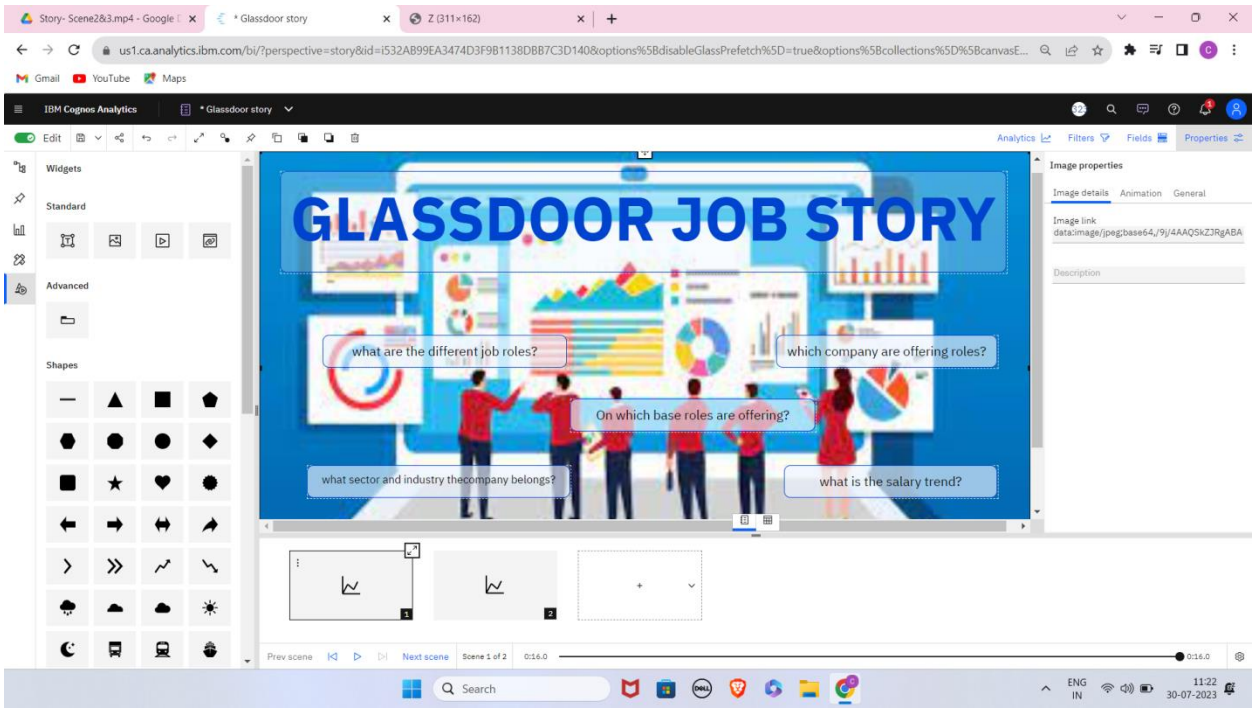

Story- Scene2&3.mp4 - Google
* Glassdoor story
Z (311×162)
us1.ca.analytics.ibm.com/bi/?perspective=story&id=i532AB99EA3474D3F9B1138DBB7C3D140&options%5BdisableGlassPrefetch%5D=true&options%5Bcollections%5D%5BcanvasE...
Gmail
YouTube
Maps
IBM Cognos Analytics
* Glassdoor story
Edit
Analytics
Filters
Fields
Properties
Widgets
Standard
Advanced
Shapes
GLASSDOOR JOB STORY
what are the different job roles?
which company are offering roles?
On which base roles are offering?
what sector and industry thecompany belongs?
what is the salary trend?
Image properties
Image details
Animation
General
Image link
Description
Prev scene
Next scene
Scene 1 of 2
0:16.0
Search
ENG
IN
11:22
30-07-2023

FIG:06

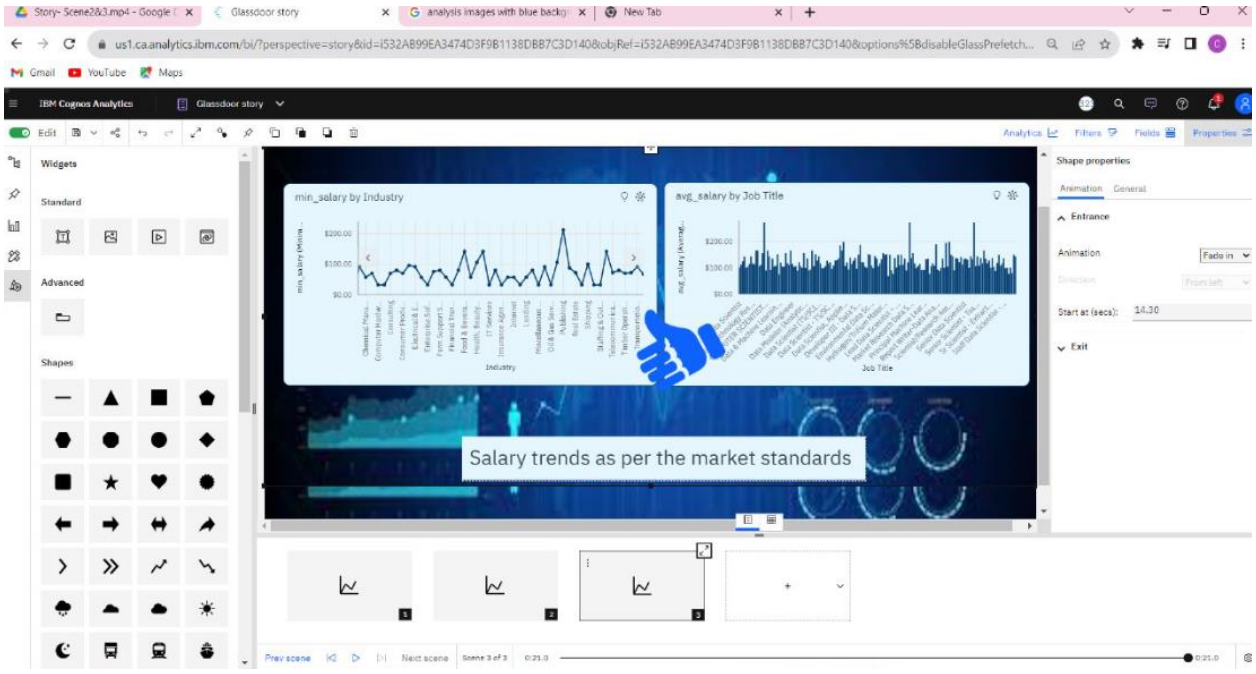

min_salary by Industry
avg_salary by Job Title
Salary trends as per the market standards

FIG:07

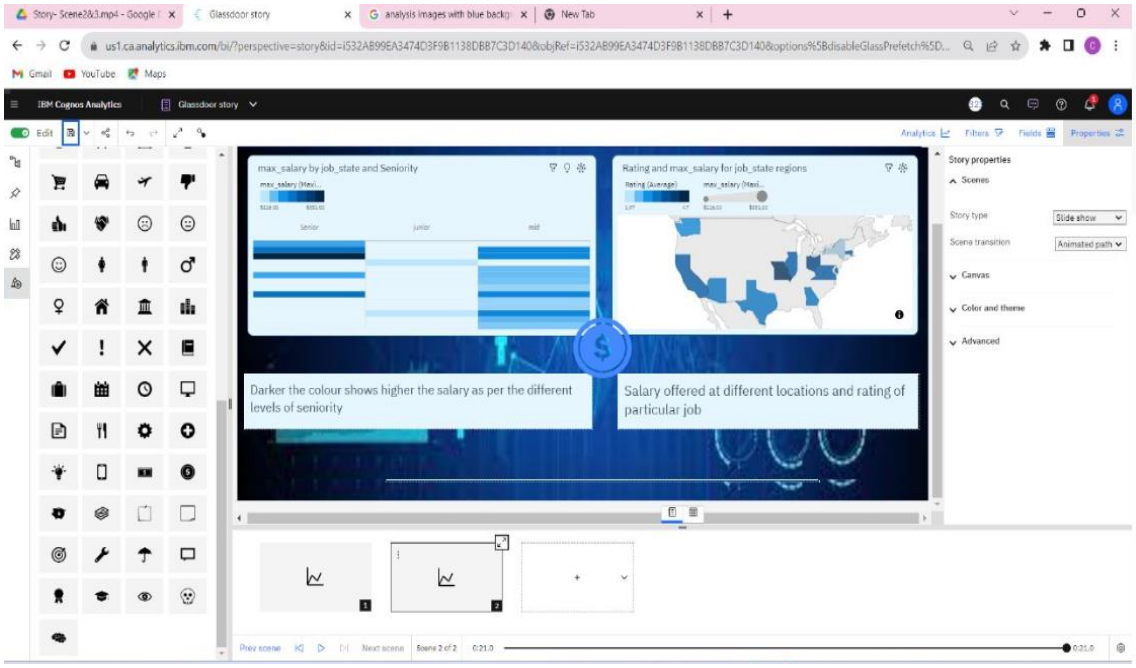

max_salary by job_state and Seniority
Rating and max_salary for job_state regions
Darker the colour shows higher the salary as per the different levels of seniority
Salary offered at different locations and rating of particular job

FIG:08


Glassdoor jobs Data Report
Which state is offering higest salary in particular domain
max_salary
$97.00
$331.00
job_simp
data scientist
analyst
na
data engineer
manager
mle
director
job_state
MA
NE
IN
MO
IL
SC
AL
NC
TX
VA
TN
OH
NJ
CT
AZ
KS
US
OK
OR
DE
Salary range for different job categories
job_simp
analyst
mle
manager
data engineer
director
data scientist
na
$350.00
$300.00
$250.00
$200.00
$150.00
$100.00
$50.00
$0.00
max_salary
analyst
mle
manager
data engineer
director
data scientist
na
job_simp

FIG:09

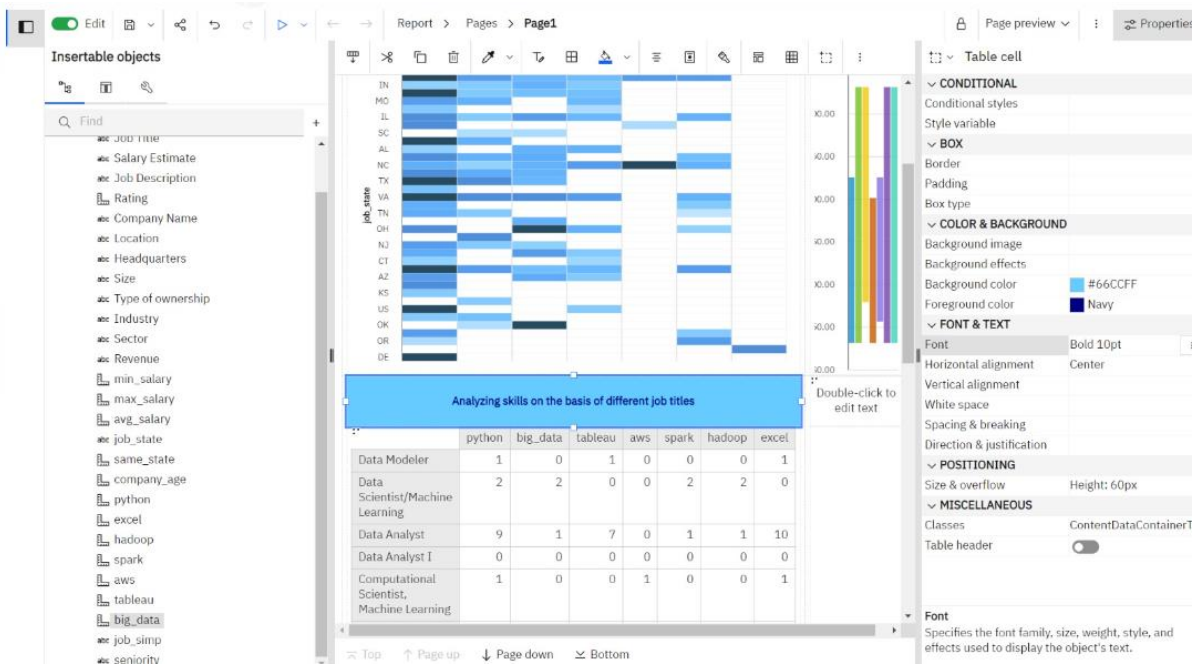

Edit
Report > Pages > Page1
Page preview
Properties
Insertable objects
Find
Salary Estimate
Job Description
Rating
Company Name
Location
Headquarters
Size
Type of ownership
Industry
Sector
Revenue
min_salary
max_salary
avg_salary
job_state
same_state
company_age
python
excel
hadoop
spark
aws
tableau
big_data
job_simp
seniority
Analyzing skills on the basis of different job titles
Double-click to edit text
python
big_data
tableau
aws
spark
hadoop
excel
Data Modeler 1 0 1 0 0 0 1
Data Scientist/Machine Learning 2 2 0 0 2 2 0
Data Analyst 9 1 7 0 1 1 10
Data Analyst I 0 0 0 0 0 0 0
Computational Scientist, Machine Learning 1 0 0 1 0 0 1
Top
Page up
Page down
Bottom
Table cell
CONDITIONAL
Conditional styles
Style variable
BOX
Border
Padding
Box type
COLOR & BACKGROUND
Background image
Background effects
Background color #66CCFF
Foreground color Navy
FONT & TEXT
Font Bold 10pt
Horizontal alignment Center
Vertical alignment
White space
Spacing & breaking
Direction & justification
POSITIONING
Size & overflow Height: 60px
MISCELLANEOUS
Classes
Table header
Font
Specifies the font family, size, weight, style, and effects used to display the object's text.

FIG:10

Analyzing skills on the basis of different job titles

	python	big_data	tableau	aws	spark	hadoop	excel
Data Modeler	1	0	1	0	0	0	1
Data Scientist/Machine Learning	2	2	0	0	2	2	0
Data Analyst	9	1	7	0	1	1	10
Data Analyst I	0	0	0	0	0	0	0
Computational Scientist, Machine Learning	1	0	0	1	0	0	1
Product Data Scientist - Ads Data Science	2	0	0	0	0	0	0
Data Scientist - Intermediate	2	0	0	0	0	0	2
Global Data Analyst	0	0	1	0	0	0	1
Data & Machine Learning Scientist	2	0	0	2	0	0	0
Data Engineer (Remote)	1	0	0	1	1	1	1
Data Scientist, Applied Machine Learning - Bay Area	2	0	0	2	0	0	0
Purification Scientist	0	0	0	0	0	0	0
Data Scientist 3 (718)	2	0	0	2	0	0	0
Real World Science, Data Scientist	2	0	0	0	0	2	0
Data Scientist - Image and Video Analytics	2	0	0	2	2	2	0
Staff Data Scientist	0	0	0	0	0	0	0
RFP Data Analyst	0	0	0	0	1	0	1
Data Scientist (TS/SCI)	1	1	0	0	0	0	0
Data Integration and Modeling	0	0	0	0	0	0	1

Max salary offered by companies at particular location in accordance with company age

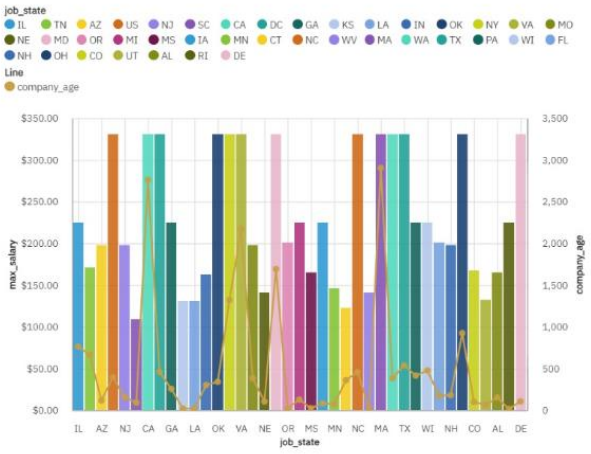


FIG:11

Github:

https://github.com/KYRLRamya/The-Future-of-Work-Data-Analysis-of-Glassdoor-jobs

5. ADVANTAGES :

The future of work data analysis from the Glassdoor jobs project can offer several advantages:

- **Insights into Labor Market Trends:** Analyzing data on job postings, salaries, and industry trends can provide valuable insights into the current and future state of the labor market, helping employers and job seekers make informed decisions.
- **Salary Transparency**: Glassdoor is known for its salary transparency feature, which benefits job seekers by allowing them to access salary information for different roles and companies, aiding in negotiations and creating a fairer job market.
- **Skill Gap Identification**: The analysis of job requirements and candidate qualifications can help identify skill gaps in the workforce, enabling educators and policymakers to address these gaps through targeted training and education programs.
- **Recruitment and Talent Acquisition**: Employers can leverage data analysis to optimize their recruitment strategies, targeting specific talent pools and streamlining the hiring process.

DISADVANTAGES:

The future of work data analysis from the Glassdoor jobs project can offer several disadvantages:

- **Data Bias and Accuracy:** Analyzing large datasets can be susceptible to biases and inaccuracies, potentially leading to skewed conclusions and misinterpretations of the job market.
- **Privacy Concerns:** Gathering and analyzing employment data on such a large scale raises privacy concerns for job seekers and employees, as sensitive information could be exposed or misused.
- **Limited Scope:** The analysis might not cover every aspect of the job market, and certain industries or regions may be underrepresented, limiting the overall accuracy and usefulness of the findings.
- **Predictive Challenges:** While data analysis can provide valuable insights, predicting future job market trends accurately remains a complex challenge due to the dynamic and ever-changing nature of the economy.

6. APPLICATIONS:

Generally, data analysis in the context of the future of work on Glassdoor could have several applications:

- **Trends in Job Market**: Analyzing job postings data to identify emerging job roles, industries, or skills in demand, helping job seekers understand the job market's evolving landscape.
- **Salary and Compensation Insights**: Providing detailed salary information and compensation trends to assist job seekers in negotiating better offers and employers in staying competitive.
- **Skill Mapping and Upskilling**: Identifying the skills employers seek the most and helping employees understand which skills to develop for career growth.
- **Diversity and Inclusion**: Using data analysis to highlight diversity and inclusion trends within companies and industries, promoting transparency and awareness.
- **Employee Satisfaction Analysis**: Leveraging employee reviews and sentiment analysis to gauge overall job satisfaction and workplace experiences within different companies.
- **Remote Work Trends**: Analyzing job postings to understand the prevalence and characteristics of remote work opportunities.
- **Job Market Predictions**: Using historical data to make predictions about future job demand and industry growth.

Remember that these are potential applications based on the general scope of Glassdoor's data. For the most up-to-date and accurate information, it's best to refer directly to Glassdoor's official website or press releases for any recent developments.

7. CONCLUSION:

The future work for the data analysis of the Glassdoor jobs project holds promising opportunities for further exploration and improvement. With the wealth of data available on the platform, there are several areas that can be focused on to enhance the analysis and provide more valuable insights.

Firstly, incorporating machine learning algorithms and natural language processing techniques can help in extracting deeper insights from the job descriptions and reviews. By analyzing the language used and identifying patterns, it is possible to gain a better understanding of the skills and qualifications required for different positions, as well as the overall sentiment associated with certain companies or industries.

Secondly, expanding the scope of the analysis beyond job postings and reviews to include other relevant factors such as salaries, benefits, and company culture can provide a more comprehensive view of the job market. This can involve incorporating data from other sources or integrating with APIs to gather additional information.

Furthermore, conducting comparative analysis between different industries, regions, or job roles can help in identifying trends and patterns, as well as highlighting areas of strength or areas for improvement. This can assist job seekers in making informed decisions and companies in benchmarking their offerings against industry standards.

Additionally, visualizing the data in meaningful and interactive ways can improve the usability and accessibility of the insights generated from the analysis. This can involve creating dashboards, interactive charts, or heat maps to enable users to explore the data and gain insights at a glance.

Lastly, continuously updating and maintaining the data analysis pipeline will be crucial for ensuring the accuracy and relevance of the insights. This includes regularly scraping new job postings and reviews from Glassdoor, cleaning and preprocessing the data, and re-running the analysis periodically to capture any changes or updates in the job market.

Overall, the future work for the data analysis of the Glassdoor jobs project holds immense potential in terms of uncovering valuable insights, providing actionable recommendations, and empowering job seekers and companies alike. By leveraging advanced analytical techniques and continuously improving the data analysis pipeline, the project can contribute to enhancing transparency and efficiency in the job market.

8. FUTURE SCOPE:

The future scope of the data analysis of Glassdoor jobs holds immense potential and can be pursued in several directions. Here are some key areas that can be focused on:

- Predictive Analytics
- Recommender Systems
- Sentiment Analysis
- Diversity and Inclusion Analysis
- Skill Gap Analysis
- Geographical Analysis
- Social and Network Analysis

Overall, the future work of data analysis of Glassdoor jobs offers numerous possibilities for enhancing the job search experience, improving transparency in the job market, promoting diversity and inclusion, and enabling data-driven decision-making for both job seekers and employers. By exploring these areas, it is possible to unlock valuable insights that can transform the way people find and evaluate job opportunities.