



Humanlytics

Feb 19, 2024

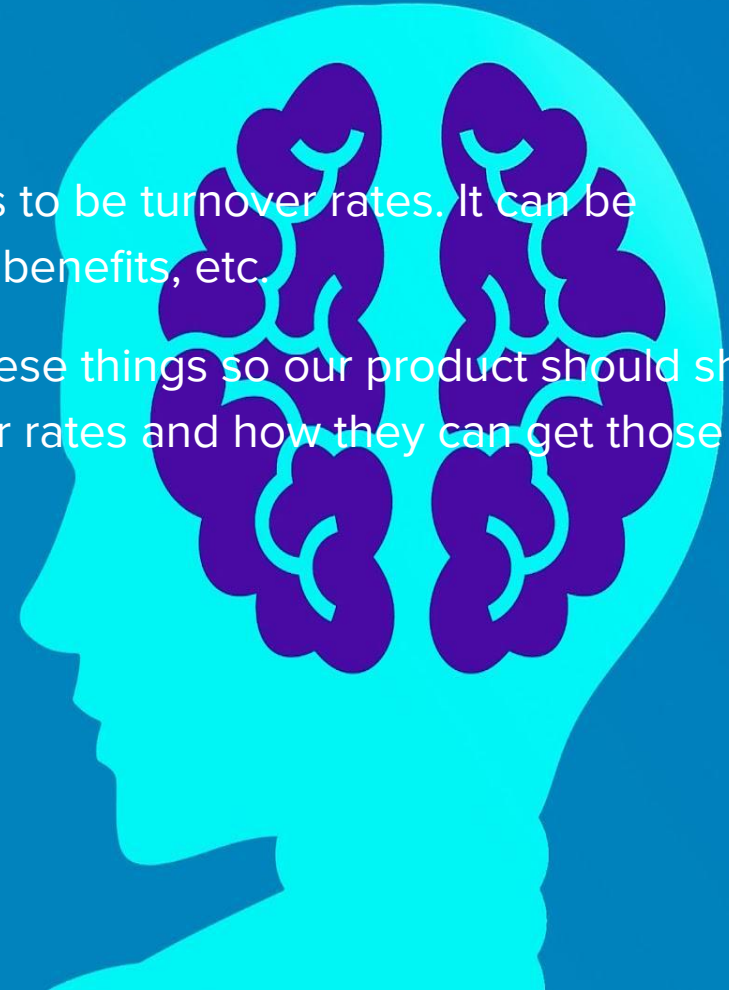
Muhammad Tayyab
Ramya
Dallas Diaz
Muhammad Arslan

https://github.com/mtayyab25/Capstone_CS-5588-
[https://github.com/THAMBABATTULARAMYAKUMARI](https://github.com/THAMBABATTULARAMYAKUMARI2019/Datascience-capstone_CS5)
[2019/Datascience-capstone_CS5](https://github.com/dallasdiaz93/datasciencecapstone)
<https://github.com/dallasdiaz93/datasciencecapstone>

Introduction

The main issue in today's work world seems to be turnover rates. It can be attributed to COVID, employee satisfaction, benefits, etc.

Not all HR teams within companies know these things so our product should shed more light on what is affecting their turnover rates and how they can get those down

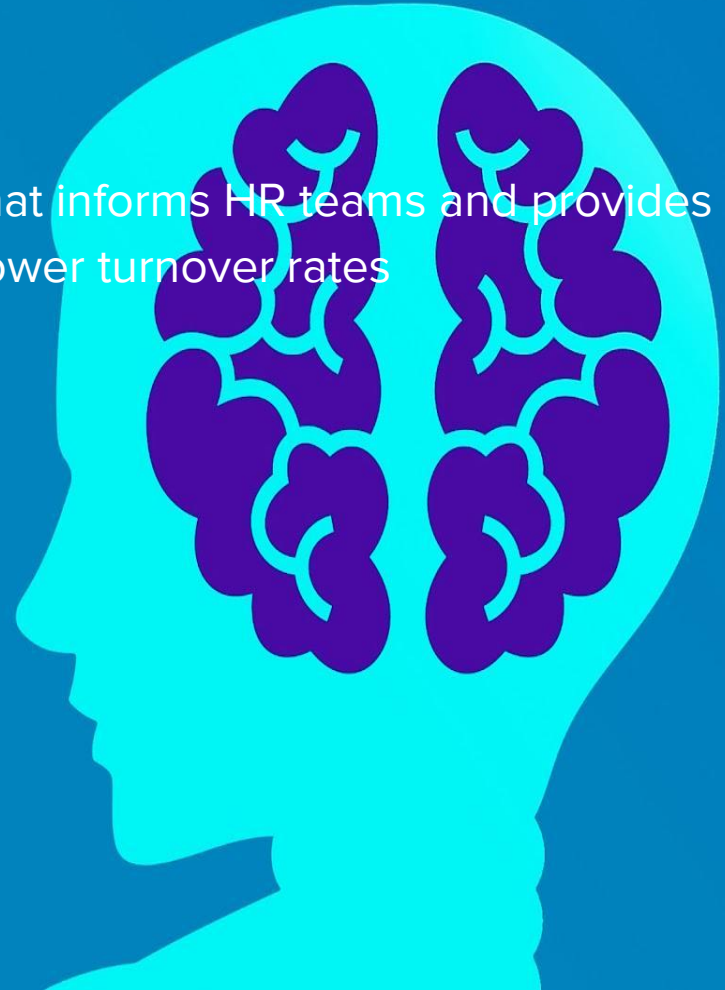


Objectives

Our Objectives are to create a dashboard that informs HR teams and provides recommendations on how to improve and lower turnover rates

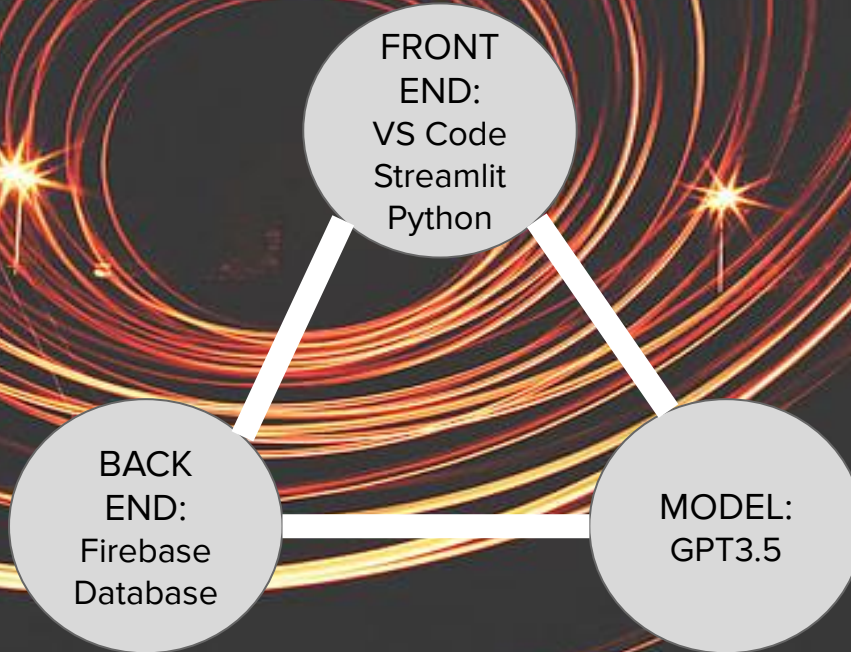
We will look at:

- Employee satisfaction
- Skills required for new hires
- Recommendations for HR staff



Triangle Model Overview

Here are our three main components of our Triangle Model:



Backend Functionality

Our main storage of static data and real time data is Firebase.

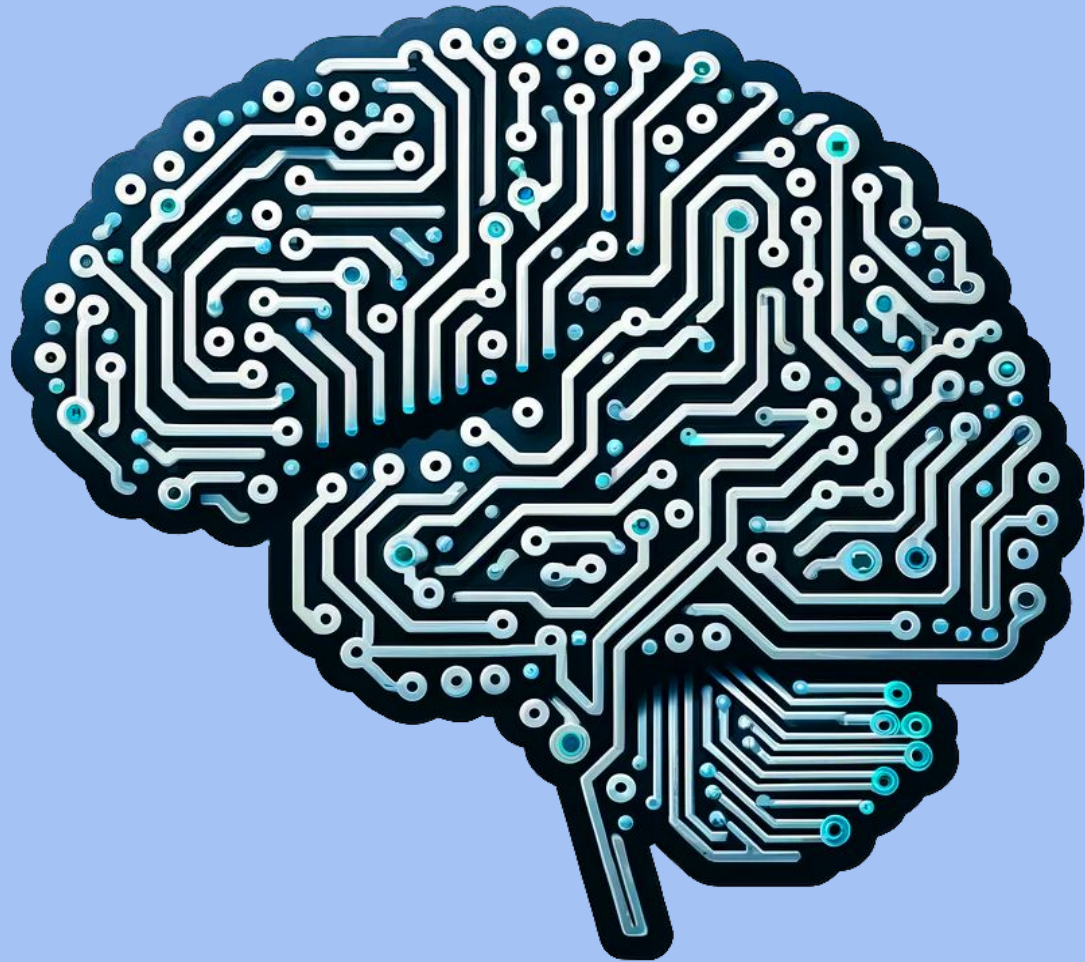
- We are inputting data sets that we've found into tables created in Firebase
- Our front end allows users to input information that goes to those same tables.



AI ANALYTICS

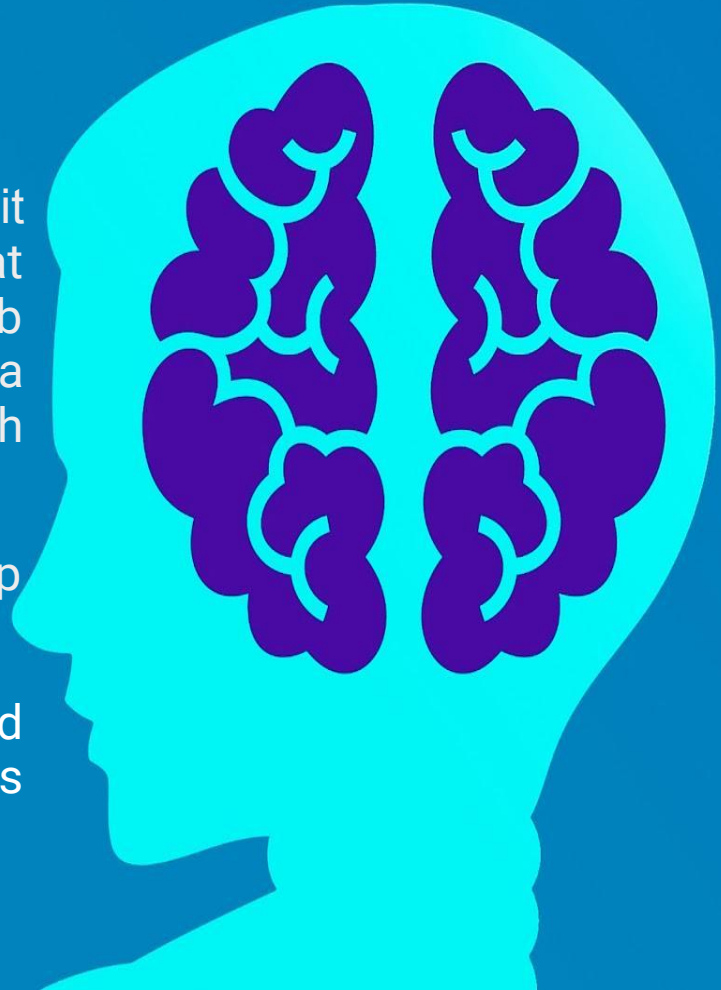
AI model Being used in this project

- GPT3.5 model
- VS code
- Streamlit
- Jupyter notebook

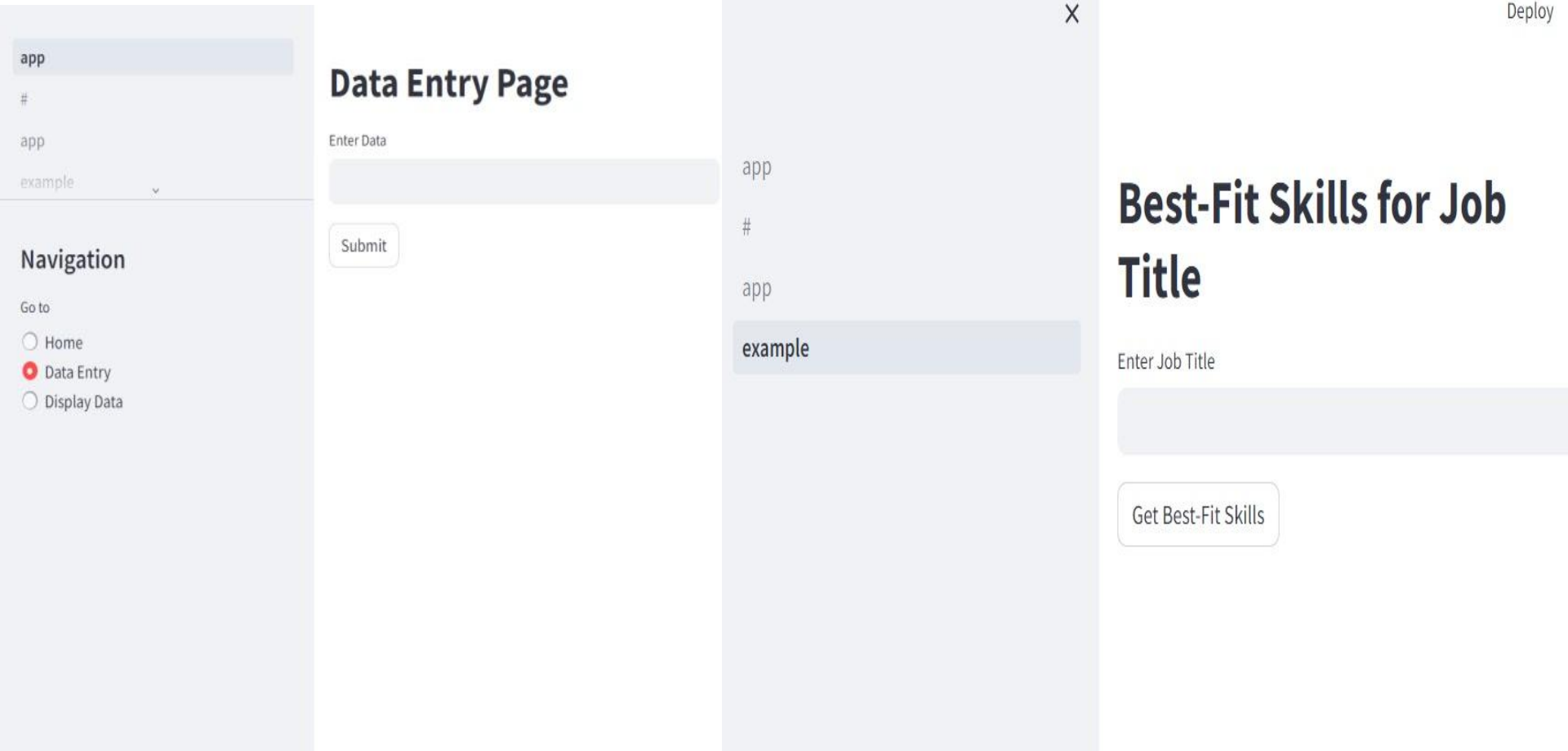


Frontend/User Interface

- User Interface is designed using Streamlit is an open-source Python library that allows developers to create web applications for machine learning, data science, and other analytical tasks with minimal effort.
- Here , Features we have developed App and Example.
- Data entry , Entering job titles into the field will prompt the display of necessary skills associated with that position .
-



User Interface



Raw Data

	A	B	C	D	E	F	G	H	I		A	B	C	D	E	F	G	H	I	J	K		
1	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raise	1	user_name	user_location	user_description	user_created	user_followers	user_friends	user_favorites	user_verified	date	text	hashtags	source	
2	New Work	Hamburg	Consumer	400		2024-01-11	Post-IPO	Germany															
3	Playtika	Tel Aviv	Consumer	300	0.1	2024-01-11	Post-IPO	Israel		2										Apply today! https://t.co/eCRVPFE			
4	Discord	SF Bay Area	Consumer	170	0.17	2024-01-11	Series H	United States			BestCompanies2	Arizona	We validate and elev	2009-01-11 16:28:33	2836	2469	870	FALSE	2022-12-20 18:00:46	#apply #hiring #open [apply', 'hiring', 'open		Hot	
5	Inmobi	Bengaluru	Marketing	125	0.05	2024-01-11	Unknown	India		3										Apply today! https://t.co/eCRVPFE			
6	Audible	New York City	Media	100	0.05	2024-01-11	Acquired	United States			BestCompanies2	Arizona	We validate and elev	2009-01-11 16:28:33	2835	2469	870	FALSE	2022-12-20 18:00:46	#apply #hiring #open [apply', 'hiring', 'open		Hot	
7	Sisense	New York City	Data	60	0.13	2024-01-11	Series F	United States		4	CHILUDO		AMC APE MAKING / 2022-03-27 17:10:56		19	32	4	FALSE	2022-12-18 00:21:02	I want to wish all @Twitter employees that Twit			
8	Google	SF Bay Area	Consumer	1000		2024-01-10	Post-IPO	United States		5	ANIL	New Delhi, India, Frc	New Delhi. Above pc	2010-04-17 16:58:14		81	287	4484	FALSE	2022-12-17 14:04:44	@elonmusk sir has cr [TwitterLayoffs]		Twit
9	Beam Benefits	Columbus	Healthcare	74		2024-01-10	Series E	United States		6	My Head's Explodin	leswarden@masodo	A Proud Believer in	2009-01-28 01:03:16		1044	777	3742	FALSE	2022-12-16 19:47:56	Here's what happens [TicKtOk]		Twit
10	Instagram	SF Bay Area	Consumer	60		2024-01-10	Acquired	United States		7	Luke Petersen		33years+ born in the	2012-01-21 09:43:13		19	105	5833	FALSE	2022-12-16 09:18:57	@Ourisid15 @Adam [Elon]		Twit
11	Amazon	Seattle	Retail			2024-01-10	Post-IPO	United States												#US Judge Asks @TV			
12	LargePoint	SF Bay Area	Manufacturing		0.12	2024-01-10	Unknown	United States												Read: https://t.co/93'			
13	Citrix	Miami	Infrastructure		0.12	2024-01-10	Acquired	United States			BW Businessworld	New Delhi	BW Businessworld is	2009-05-01 05:57:27		21539	437	5272	FALSE	2022-12-16 07:58:00	#twitter #ayoffs#twil [US]		Two
14	SoFi	SF Bay Area	Finance		0.04	2024-01-10	Post-IPO	United States		9	Paze+Amor	NoCal	ELL teacher 30+ yrs	2010-07-17 23:36:36		395	1186	106249	FALSE	2022-12-16 01:13:16	@wamamubell he ew [TwitterLayoffs]		Twit
15	Twitch	SF Bay Area	Consumer	500	0.35	2024-01-09	Acquired	United States		10										We just filed yet ano			
16	Branch	Columbus	Finance	85		2024-01-09	Series C	United States			Lisa Bloom	Los Angeles	Trial lawyer fighting	2008-08-26 14:09:47		133631	952	3481	TRUE	2022-12-16 00:51:02	#TwitterLayoffs [TwitterLayoffs]		Twit
17	Nevro	SF Bay Area	Healthcare	63	0.05	2024-01-09	Post-IPO	United States		11	The White Rabbit	Naples, Italy	U.S. Navy Vet. No Pr	2022-11-26 22:33:27		276	526	1731	FALSE	2022-12-15 17:17:44	We got sickup kids, corrupt cops, and cra		
18	Uber Freight	SF Bay Area	Logistics	40		2024-01-09	Subsidiary	United States		12	KevinB	Arizona	Internet made simple	2022-11-28 20:00:16		6	54	4	FALSE	2022-12-15 01:56:22	@elonmusk Hey papa elon @elonmusk, w [TwitterLayoffs] #hisshouldn'tbe legal https://		
19	Rent the Runway	New York City	Retail	37	0.1	2024-01-09	Post-IPO	United States												I own a new social communication app Now on Android and iOS Let's start a new company and launch the #TwitterLayoffs #TwitchStreamers #TwitchAffiliate #smallstreamer			
20	Humane	SF Bay Area	Hardware	10	0.04	2024-01-09	Series C	United States		13													
21	Trend Micro	Tokyo	Security		0.02	2024-01-09	Unknown	Japan			Abdullah	San Francisco, CA	CEO T one Social M	2021-11-27 18:19:16		27	423	12	FALSE	2022-12-14 18:51:07	#elonsoffiliat		Twit
22	Unity	SF Bay Area	Other	1800	0.25	2024-01-08	Post-IPO	United States		14	Mario Di Maggio	Bristol	Senior science com	2021-09-10 20:00:12		98	50	3530	FALSE	2022-12-14 17:33:25	Em @elonmusk, acc [LoveWhereYouWor		Twit
23	NuScale Power	Corvallis	Energy	154	0.28	2024-01-08	Post-IPO	United States															
24	Pitch	Berlin	Other	80	0.67	2024-01-08	Series B	Germany															
25	BenchSci	Toronto	Healthcare	70	0.17	2024-01-08	Series D	Canada															
26	Flexe	Seattle	Logistics		0.38	2024-01-08	Series D	United States													The lawsuit by forme		
27	Here	Miami	Real Estate		1	2024-01-08	Seed	United States		15											Understand your #ig		

Data Cleaning and PreProcessing

Data Cleaning

```
[5] # Handle missing values
df.fillna({'percentage_laid_off': 0}, inplace=True)
```

df



	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raised
0	New Work	Hamburg	Consumer	400.0	0.00	2024-01-11	Post-IPO	Germany	NaN
1	Playtika	Tel Aviv	Consumer	300.0	0.10	2024-01-11	Post-IPO	Israel	NaN
2	Discord	SF Bay Area	Consumer	170.0	0.17	2024-01-11	Series H	United States	995.0
3	Inmobi	Bengaluru	Marketing	125.0	0.05	2024-01-11	Unknown	India	320.0
4	Audible	New York City	Media	100.0	0.05	2024-01-11	Acquired	United States	14.0
...
3308	Service	Los Angeles	Travel	NaN	1.00	2020-03-16	Seed	United States	5.1
3309	HopSkipDrive	Los Angeles	Transportation	8.0	0.10	2020-03-13	Unknown	United States	45.0
3310	Panda Squad	SF Bay Area	Consumer	6.0	0.75	2020-03-13	Seed	United States	1.0
3311	Tamara Mellon	Los Angeles	Retail	20.0	0.40	2020-03-12	Series C	United States	90.0
3312	EasyPost	Salt Lake City	Logistics	75.0	0.00	2020-03-11	Series A	United States	12.0

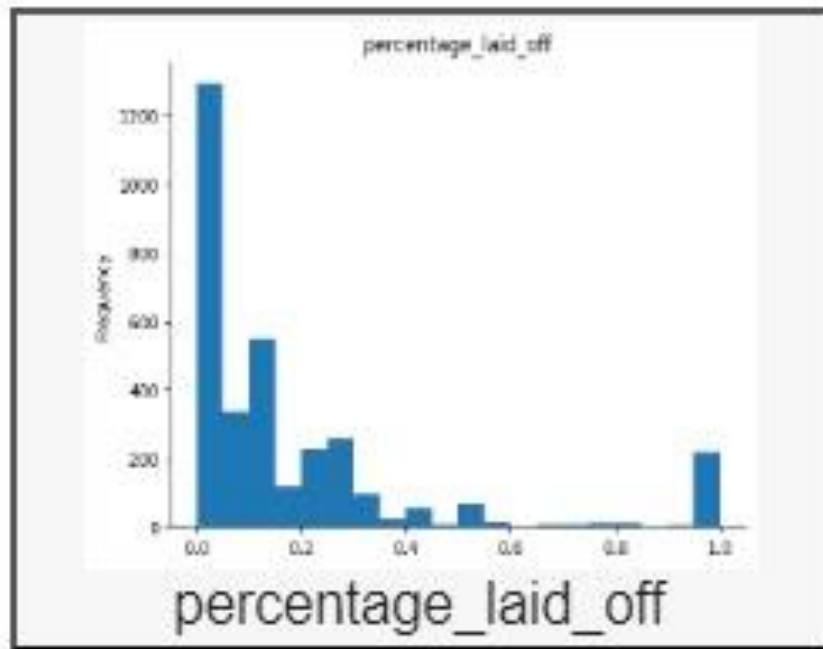
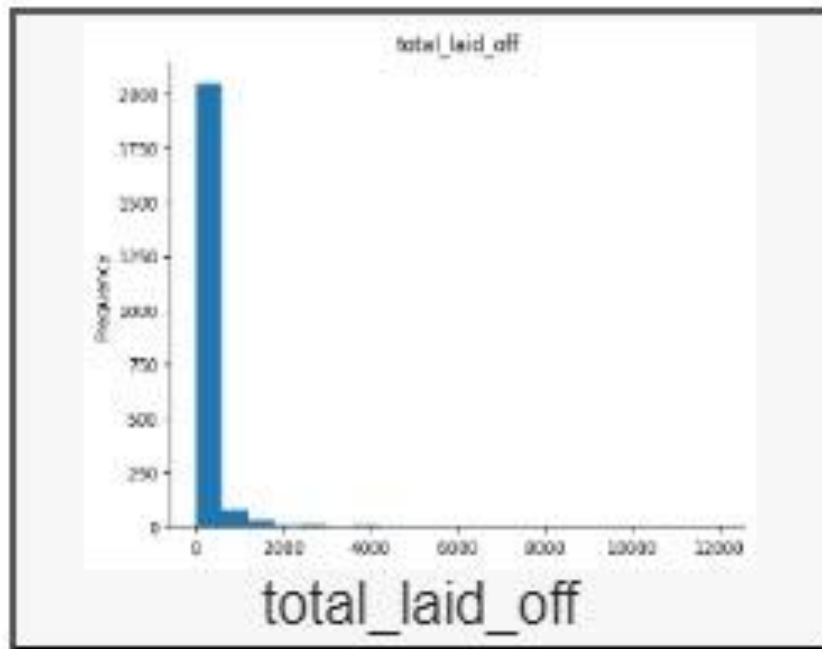
3313 rows × 9 columns

Firestore Connection

```
capstone_project
> __pycache__
✓ content
  processed_data.csv
  Screenshot 2023-12-13 135...
> key
✓ pages
  #.py
  app.py
  example.py
  ml.ipynb
  app.py

12 # Define functions to interact with Firestore
13 def write_data(data):
14     db.collection('data').add(data)
15
16 def read_data():
17     docs = db.collection('skills').stream()
18     data_list = [doc.to_dict() for doc in docs]
19     return data_list
20
21 # Define Streamlit pages
22 def home():
23     st.title("Home Page")
24     # Add content for home page
25
26 def data_entry():
27     st.title("Data Entry Page")
28     # Add form for data entry
29     # Example:
30     data = st.text_input("Enter Data")
31     if st.button("Submit"):
32         write_data({"data": data})
33
34 def display_data():
35     st.title("Display Data Page")
36     # Fetch and display data from Firestore
37     data_list = read_data()
38     for data in data_list:
39         st.write(data)
40
```

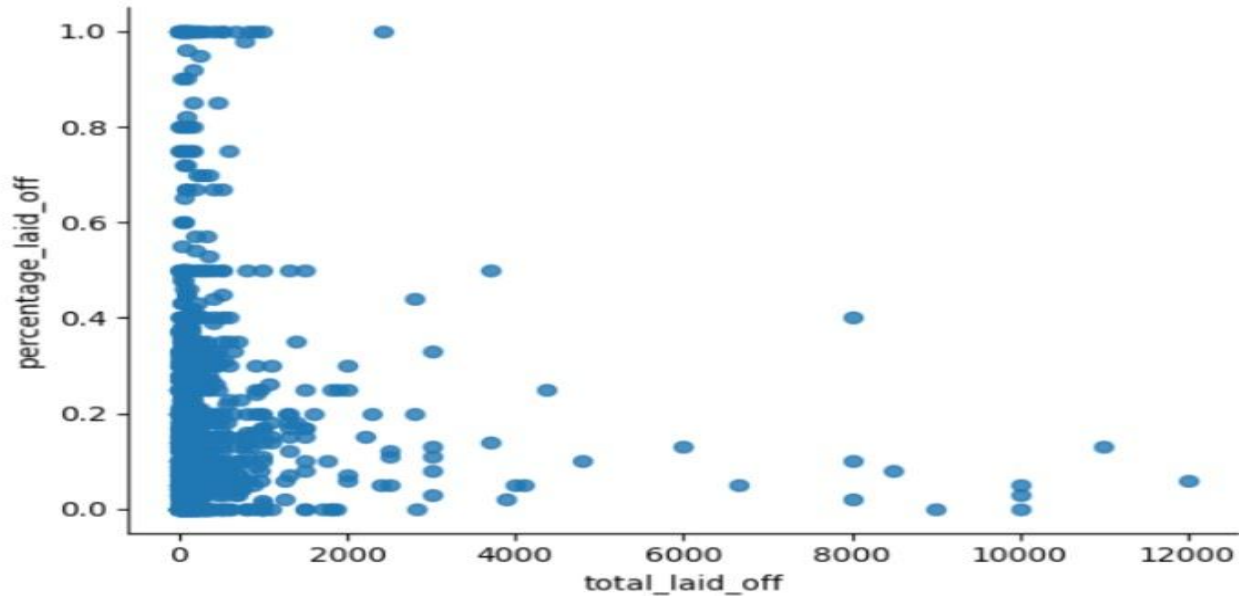
Layoff data_Visualization



Layoff_Visualization



```
from matplotlib import pyplot as plt
df.plot(kind='scatter', x='total_laid_off', y='percentage_laid_off', s=32, alpha=.8)
plt.gca().spines[['top', 'right',]].set_visible(False)
```



Integration

- The integration of GPT-3.5 model, VS Code, and Streamlit forms the foundation of our dashboard. GPT-3.5 model provides natural language processing capabilities to interpret user inputs and generate recommendations. VS Code serves as the development environment for coding and managing the project. Streamlit acts as the user interface framework, enabling seamless interaction and visualization of data and recommendations.

Integration Challenges and Solutions

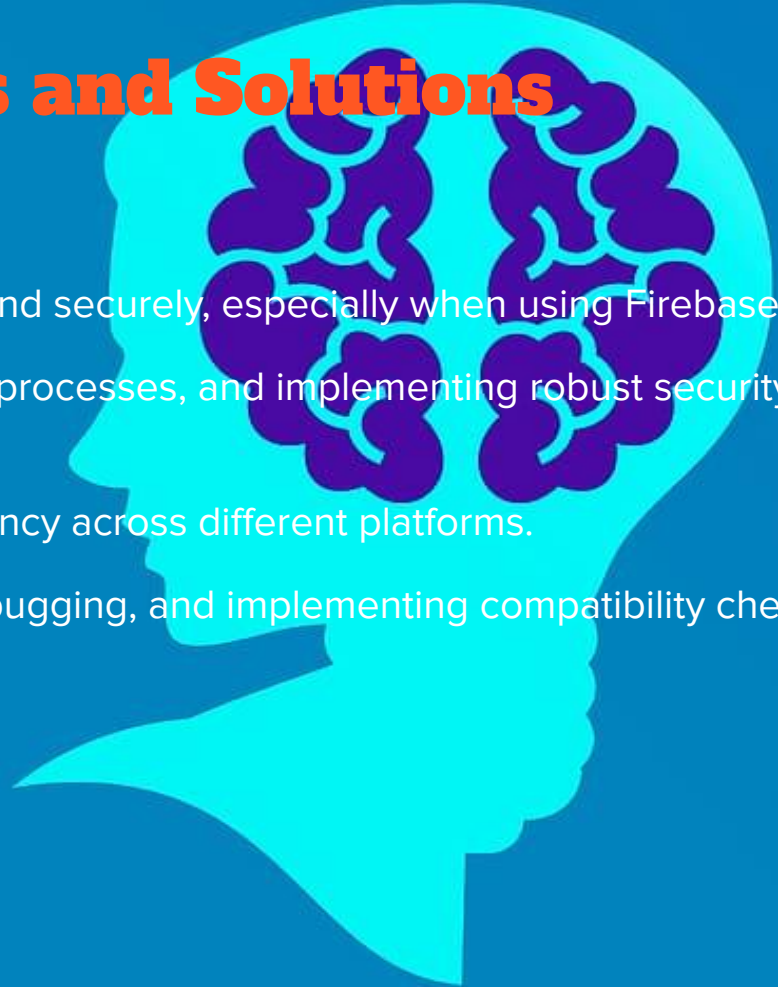
Here are Challenges and

Challenge 1: Handling large datasets efficiently and securely, especially when using Firebase.

Solution 1: Optimizing data storage and retrieval processes, and implementing robust security measures to protect sensitive information.

Challenge 2: Ensuring compatibility and consistency across different platforms.

Solution 2: Conducting thorough testing and debugging, and implementing compatibility checks and version control mechanisms.



Contribution Points For Integration Efforts

- Dallas: 25%
- Arslan: 25%
- Tayyab: 25%
- Ramya: 25%

Results and Discussion

- The dashboard informs HR teams of turnover rate contributors like employee satisfaction and necessary skills, enabling data-driven decisions. This empowers HR to tackle issues effectively and implement retention strategies.
- The Streamlit-designed interface provides HR teams with a seamless, intuitive experience, allowing easy data input and visualization. Integration of AI models like GPT-3.5 enhances usability, making it a valuable tool for addressing turnover challenges.

Data Collection

- https://docs.google.com/spreadsheets/d/13LGiJ2aAFWkWvqfayr3bSWIJOvvx4qNsJ_6RFvAUXb0/edit#gid=1208650237
- <https://docs.google.com/spreadsheets/d/1TZc5gju6TXprXaLimRwyrWhF4umrBEpun4TuV5xcIIA/edit#gid=1508777424>
- https://docs.google.com/spreadsheets/d/1M_2IrocTETVPCIWdZStjdfTqTROA7aJf8Y0WqmdG6KQ/edit#gid=1513922754
- <https://docs.google.com/spreadsheets/d/1kwamdTZ8bphHfhDrEdiNA5so2rYiWY0/edit#gid=88241723>
- <https://drive.google.com/drive/search>

Conclusion

Summary of Achievements: We've combined Firebase for storing data, GPT-3.5 for AI analysis, and Streamlit for creating the dashboard. This makes our tool complete for tackling turnover rates at work, giving HR teams useful insights and suggestions to handle turnover better.

Impact of Integrated Components: Bringing together Firebase, GPT-3.5, and Streamlit has greatly improved our dashboard usability. HR teams can now understand turnover factors, get new hire skill recommendations, and make better decisions for improving outcomes.

Future Work Suggestions: We plan to make our dashboard even better by adding more data sources and improving our AI algorithms. We also aim to expand its use beyond turnover rates to help HR teams with other workforce management challenges, ensuring continuous improvement and value.

Reference

- [1] https://github.com/r Rambhia22/ResumeClassification_Parser
- [2] <https://github.com/ghelanikirtan/HR-Process-Automation>
- [3] <https://github.com/deepak147/resumeClassify>
- [4] <https://gist.github.com/cquangc/1b0d6cf08fe250a680d1d823eff>
- [5] https://github.com/Spidy20/Smart_Resume_Analyser_App

Thank you for you time

Questions? Please feel free to contact

Dallas Diaz: ddmnh@umsystem.edu

Muhammad Tayyab : mt56w@umsystem.edu

Ramya Thambabattula: tramyakumari.22s@gmail.com

Muhammad Arslan: manbb@umsystem.edu