People Analytics Project

Madison Marchionna

For this project, I have chosen to create visualizations using Python. The size of the dataset allowed me to do so quickly and efficiently using three Pandas Dataframes. However, to demonstrate my ability to write SQL queries, I have written the following code that could be used to analyze the data in a RDBMS (I used PostgreSQL).

Overall Data Interpretation:

* Layoffs associated with COVID-19 significantly influenced total turnover rate in 2020.
* April is the highest turnover month, across reporting periods. Additionally, we see higher turnover in the summer months of July and August.
* Outside of involuntary attrition, more than 50% of the turnover is regrettable.
* G&A turnover is consistently higher than that of R&D and Sales & Marketing.
* Turnover is trending positive in 2022 as compared to prior years.
* Turnover in the Americas is consistently lower than that of other regions.

**Query #1:**

select \* from formlabsproject

where total\_attrition > 19

Results:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Date | Headcount | Total Attrition | Regretted Attrition | Non-Regretted Attrition | Involuntary Attrition |
| "2020-01-01" | 573 | 20 | 11 | 6 | 3 |
| "2020-04-01" | 531 | 27 | 2 | 1 | 24 |
| "2021-06-01" | 615 | 22 | 14 | 7 | 1 |
| "2021-07-01" | 616 | 24 | 16 | 5 | 3 |
| "2021-08-01" | 637 | 24 | 12 | 10 | 2 |

This table shows the top five months with the greatest attrition counts.

**Query #2:**

SELECT

DATE\_TRUNC('quarter', date) AS quarter,

SUM(regretted\_attrition + non\_regretted\_attrition + involuntary\_attrition) AS total\_attrition,

SUM(regretted\_attrition) AS regretted\_attrition,

SUM(non\_regretted\_attrition) AS voluntary\_non\_regretted\_attrition,

SUM(involuntary\_attrition) AS involuntary\_attrition

FROM

formlabsproject

GROUP BY

quarter

ORDER BY

Quarter;

Results:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date | Total Attrition | Regretted Attrition | Non-Regretted Attrition | Involuntary Attrition |
| 2018 Q3 | 28 | 14 | 10 | 4 |
| 2018 Q4 | 32 | 9 | 16 | 7 |
| 2019 Q1 | 24 | 14 | 6 | 4 |
| 2019 Q2 | 27 | 17 | 7 | 3 |
| 2019 Q3 | 42 | 29 | 10 | 3 |

…

This produces the attrition counts per quarter, which can be used to create a more concise visualization of the change over time (as opposed to using every month).

**Query #3:**

SELECT

EXTRACT(MONTH FROM date) AS month,

SUM (regretted\_attrition + non\_regretted\_attrition + involuntary\_attrition) AS total\_attrition

FROM

formlabsproject

GROUP BY

month

ORDER BY

month;

Results:

|  |  |
| --- | --- |
| Month | Total Attrition |
| 1 | 47 |
| 2 | 43 |
| 3 | 59 |
| 4 | 72 |

…

This query calculates the attrition count for each of the twelve months, which can be used to determine what time of the year has a higher/lower attrition rate.