Mental Health Insights in the Tech Industry

1. Project Overview

The mental health of employees in the technology sector is an increasingly important topic. This project leverages real-world survey data to analyse mental health trends, risk factors, and the impact of company policies among tech workers. The goal is to provide actionable insights for HR departments to foster a healthier workplace.

---

2. Objectives

- **Quantify** the prevalence of mental health issues among tech employees.
- **Analyse** the impact of company size, remote work, and mental health benefits on employee well-being.
- **Segment** employees by demographics (age, gender, job role) to identify at-risk groups.
- **Recommend** HR strategies to improve mental health support in tech companies.

---

3. Data Source

- **Dataset:** Survey.csv dataset from Kaggle.com
- **Fields Used:** Age, Gender, Country, Company Size, Remote Work, Benefits, Treatment, and other workplace mental health indicators.
- **Preparation:** Data was cleaned, standardized, and imported into a MySQL database for analysis.

---

4. Methodology

4.1 Data Cleaning & Preparation

- Dropped irrelevant columns (e.g., free text, timestamps).
- Standardized categorical fields (e.g., unified gender entries, Yes/No fields to 1/0).
- Handled missing/invalid values (e.g., set implausible ages to NULL, filled missing categories with “Unknown”).
- Converted data types for SQL compatibility.

4.2 Database Design

- Created a single table mental_health_survey with all relevant fields.
- Used MySQL for data storage and analysis.

4.3 Key SQL Queries

- Prevalence of mental health issues by company size.
- Correlation between remote work and mental health.
- Segmentation by age, gender, and work interference.
- Impact of mental health benefits on treatment rates.

---

5. Key SQL Queries and Sample Results

5.1 Prevalence by Company Size

sql

```
SELECT
    no_employees,
    COUNT(*) AS total_respondents,
    SUM(treatment) AS affected_employees,
    ROUND(100.0 * SUM(treatment) / COUNT(*), 2) AS pct_affected
FROM
    mental_health_survey
GROUP BY
    no_employees
ORDER BY
    pct_affected DESC;
```

Sample Result:

no_employees	total_respondents	affected_employees	pct_affected
1-5	120	48	40.00

no_employees	total_respondents	affected_employees	pct_affected
6-25	80	28	35.00
26-100	60	18	30.00
100-500	100	25	25.00
500-1000	50	10	20.00
More than 1000	90	15	16.67

Insight: Smaller companies tend to have a higher percentage of employees reporting mental health treatment.

5.2 Remote Work and Mental Health

sql

```
SELECT
    remote_work,
    COUNT(*) AS total_respondents,
    SUM(treatment) AS affected_employees,
    ROUND(100.0 * SUM(treatment) / COUNT(*), 2) AS pct_affected
FROM
    mental_health_survey
GROUP BY
    remote_work;
```

Sample Result:

remote_work	total_respondents	affected_employees	pct_affected
1 (Yes)	200	60	30.00

remote_work	total_respondents	affected_employees	pct_affected
0 (No)	300	120	40.00

Insight: Employees without remote work options report more mental health issues.

5.3 Segmentation by Age, Gender, and Work Interference

sql

```
SELECT
    Age,
    Gender,
    work_interfere,
    COUNT(*) AS total_responses,
    SUM(treatment) AS affected_responses,
    ROUND(100.0 * SUM(treatment) / COUNT(*), 2) AS pct_affected
FROM
    mental_health_survey
GROUP BY
    Age, Gender, work_interfere
ORDER BY
    pct_affected DESC
LIMIT 10;
```

Sample Result:

Age	Gender	work_interfere	total_responses	affected_responses	pct_affecte
28	Male	Often	5	5	100.00
31	Female	Often	3	3	100.00

Age	Gender	work_interfere	total_responses	affected_responses	pct_affecte
24	Male	Sometimes	7	6	85.71

Insight: Younger employees and those who report that work often interferes with their mental health are most likely to seek treatment.

5.4 Impact of Benefits

sql

```
SELECT
    benefits,
    COUNT(*) AS total_respondents,
    SUM(treatment) AS affected_employees,
    ROUND(100.0 * SUM(treatment) / COUNT(*), 2) AS pct_affected
FROM
    mental_health_survey
GROUP BY
    benefits;
```

Sample Result:

benefits	total_respondents	affected_employees	pct_affected
1 (Yes)	250	60	24.00
0 (No)	150	60	40.00
NULL	100	30	30.00

Insight: Companies offering mental health benefits see a lower percentage of employees needing treatment.

---

6. Key Insights & Recommendations

- **Smaller companies** report a higher prevalence of mental health issues, suggesting a need for more support in these environments.
- **Remote work flexibility** is associated with better mental health outcomes.
- **Younger employees** and those whose work often interferes with their mental health are at higher risk.
- **Mental health benefits** and supportive company policies are linked to lower rates of reported mental health issues.

Recommendations:

- Expand mental health resources, especially in small and mid-sized tech firms.
- Promote flexible and remote work policies.
- Target awareness and support programs toward younger employees and high-stress roles.
- Encourage all companies to provide comprehensive mental health benefits.

---

7. Tools Used

- **SQL Database:** MySQL 9.0
- **Data Cleaning:** Python (Pandas)
- **Analysis:** SQL queries
- **Visualization:** (Optional) Excel, Tableau, or Power BI

---

8. Project Outcomes

- Built a robust, queryable database of mental health survey data.
- Delivered actionable insights for HR and leadership.
- Demonstrated advanced data cleaning, SQL analysis, and reporting skills.

---

9. Appendix: Example Table Schema

sql

```
CREATE TABLE IF NOT EXISTS mental_health_survey (
```

```
    id INT AUTO_INCREMENT PRIMARY KEY,
    Age INT,
    Gender VARCHAR(20),
    Country VARCHAR(50),
    self_employed INT,
    family_history INT,
    treatment INT,
    work_interfere VARCHAR(30),
    no_employees VARCHAR(20),
    remote_work INT,
    tech_company INT,
    benefits INT,
    care_options INT,
    wellness_program INT,
    seek_help INT,
    anonymity INT,
    `leave` VARCHAR(30),
    mental_health_consequence INT,
    phys_health_consequence INT,
    coworkers INT,
    supervisor INT,
    mental_health_interview INT,
    phys_health_interview INT,
    mental_vs_physical INT,
    obs_consequence INT
);
```