Occupational Health and Well-Being & Its Impact on Productivity

### **SDG Problem Definition Document** [**Link to pitch deck**](https://nnajichristian24.my.canva.site/occupational-health-and-well-being-its-impact-on-productivity-pptx)[**Link to spreadsheet**](https://docs.google.com/spreadsheets/d/1asMRRAJgyLE_ZWgM0x2-jBmukcsWR3nh-R8xsksCa10/edit?usp=sharing)

**Title**: **Impact of Occupation on Health and Well-Being: A Data-Driven Analysis**

**Introduction**: This project addresses two critical Sustainable Development Goals (SDGs):

* **SDG 3: Good Health and Well-Being** – which aims to ensure healthy lives and promote well-being for all at all ages.
* **SDG 8: Decent Work and Economic Growth** – which promotes sustained, inclusive, and sustainable economic growth, full and productive employment, and decent work for all.

**Problem Statement**: Occupational health and well-being are significantly influenced by the nature of the work individuals engage in. Certain occupations, particularly those with high physical demands, exposure to hazardous materials, or stressful environments, are associated with poor health outcomes. This project seeks to understand the impact of various occupations on health and well-being, focusing on identifying work-related factors that contribute to poor health outcomes.

**Objectives**:

1. To analyze the relationship between occupation types and health outcomes.
2. To identify working conditions that are most strongly associated with poor health and well-being.
3. To explore the correlation between access to healthcare and the overall health outcomes of workers in different occupations.
4. To provide actionable insights that can inform policies and interventions aimed at improving occupational health and well-being.

**Scope**: The project will involve collecting data on employees' occupations, their working conditions, and health outcomes. It will include data analysis to identify trends and correlations that can inform recommendations for improving workplace health and safety.

**Data Collection**: The data will include information on:

* **Employees**: Demographic information, occupation type, and location.
* **Occupations**: Industry, risk level, and working conditions.
* **Health Outcomes**: Diagnoses, treatments, and health status following healthcare visits.
* **Healthcare Providers**: Information on the healthcare facilities accessed by employees.

**Expected Outcomes**:

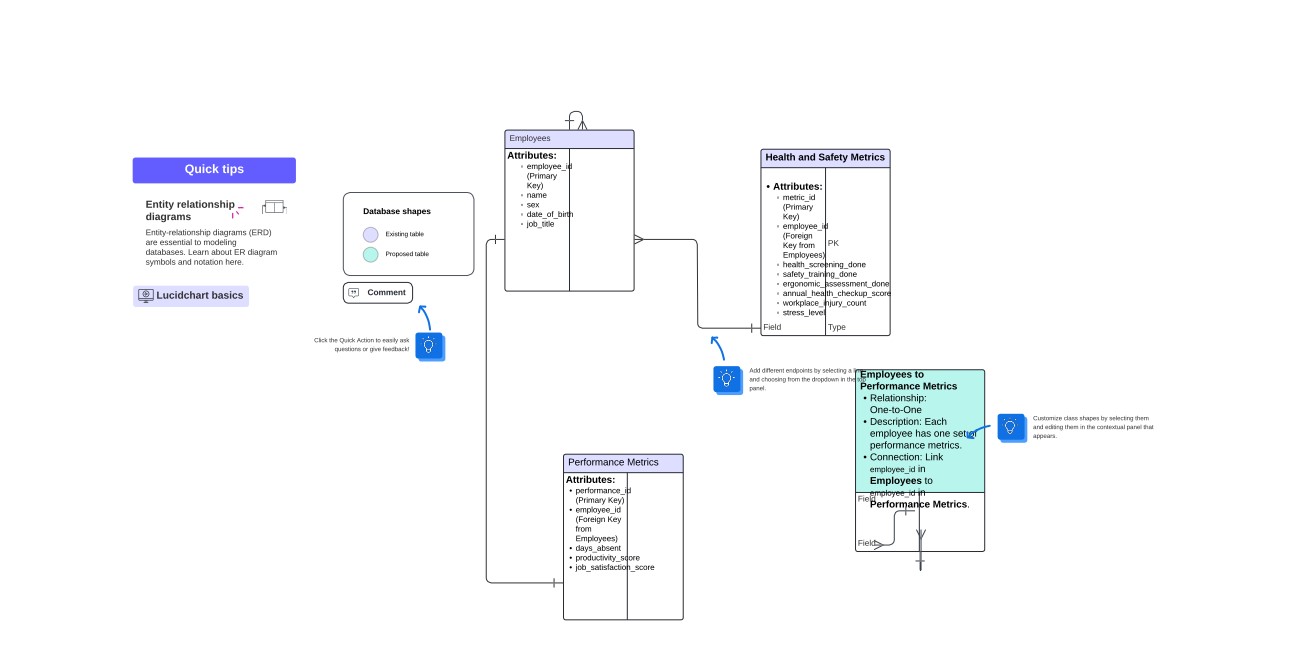
* Identification of occupations with the highest risk of poor health outcomes.
* Insights into how working conditions contribute to employee health and well-being.
* Recommendations for improving occupational health through better working conditions and increased access to healthcare.

**Conclusion**: By leveraging data on occupation and health outcomes, this project aims to contribute to the goals of ensuring good health and well-being and promoting decent work. The insights gained will be valuable for policymakers, employers, and healthcare providers working to improve the health and safety of workers across various industries.

**ERD: Design an ERD for my project, including entities relevant to my SDG problem:**

**I tried exporting my file using this data codes on Lucidchart**

SELECT 'mysql' dbms,t.TABLE\_SCHEMA,t.TABLE\_NAME,c.COLUMN\_NAME,c.ORDINAL\_POSITION,c.DATA\_TYPE,c.CHARACTER\_MAXIMUM\_LENGTH,n.CONSTRAINT\_TYPE,k.REFERENCED\_TABLE\_SCHEMA,k.REFERENCED\_TABLE\_NAME,k.REFERENCED\_COLUMN\_NAME FROM INFORMATION\_SCHEMA.TABLES t LEFT JOIN INFORMATION\_SCHEMA.COLUMNS c ON t.TABLE\_SCHEMA=c.TABLE\_SCHEMA AND t.TABLE\_NAME=c.TABLE\_NAME LEFT JOIN INFORMATION\_SCHEMA.KEY\_COLUMN\_USAGE k ON c.TABLE\_SCHEMA=k.TABLE\_SCHEMA AND c.TABLE\_NAME=k.TABLE\_NAME AND c.COLUMN\_NAME=k.COLUMN\_NAME LEFT JOIN INFORMATION\_SCHEMA.TABLE\_CONSTRAINTS n ON k.CONSTRAINT\_SCHEMA=n.CONSTRAINT\_SCHEMA AND k.CONSTRAINT\_NAME=n.CONSTRAINT\_NAME AND k.TABLE\_SCHEMA=n.TABLE\_SCHEMA AND k.TABLE\_NAME=n.TABLE\_NAME WHERE t.TABLE\_TYPE='BASE TABLE' AND t.TABLE\_SCHEMA NOT IN('INFORMATION\_SCHEMA','mysql','performance\_schema');



**Write SQL statements to create the database schema based on your ERD**

**CREATE DATABASE occupational\_health;**

**USE occupational\_health;**

CREATE DATABASE occupational\_health; USE occupational\_health; -- Create the table for employees CREATE TABLE employees ( employee\_id INT AUTO\_INCREMENT PRIMARY KEY, name VARCHAR(255), sex ENUM('Male', 'Female', 'Other'), date\_of\_birth DATE, job\_title VARCHAR(255) ); -- Create the table for health and safety metrics CREATE TABLE health\_safety\_metrics ( metric\_id INT AUTO\_INCREMENT PRIMARY KEY, employee\_id INT, health\_screening\_done BOOLEAN, safety\_training\_done BOOLEAN, ergonomic\_assessment\_done BOOLEAN, annual\_health\_checkup\_score DECIMAL(5,2), workplace\_injury\_count INT, stress\_level INT, FOREIGN KEY (employee\_id) REFERENCES employees(employee\_id) ); -- Create the table for performance metrics CREATE TABLE performance\_metrics ( performance\_id INT AUTO\_INCREMENT PRIMARY KEY, employee\_id INT, days\_absent INT, productivity\_score DECIMAL(5,2), job\_satisfaction\_score DECIMAL(5,2), FOREIGN KEY (employee\_id) REFERENCES employees(employee\_id)

### **Insert Sample Data**

INSERT INTO employees (name, sex, date\_of\_birth, job\_title)

VALUES

('Sheryl Lowery', 'Male', '1985-06-15', 'Software Engineer'),

('Sherry Caldwell', 'Female', '1990-09-22', 'Data Analyst'),

('Emily Davis', 'Female', '1987-11-05', 'Project Manager'),

('Michael Brown', 'Male', '1979-03-13', 'HR Manager'),

('Anna Wilson', 'Female', '1993-01-30', 'Sales Executive');

#### **Insert Health and Safety Metrics**

INSERT INTO health\_safety\_metrics (employee\_id, health\_screening\_done, safety\_training\_done, ergonomic\_assessment\_done, annual\_health\_checkup\_score, workplace\_injury\_count, stress\_level)

VALUES

(1, TRUE, TRUE, TRUE, 88.5, 0, 3),

(2, TRUE, TRUE, FALSE, 92.0, 1, 5),

(3, FALSE, TRUE, TRUE, 85.0, 2, 7),

(4, TRUE, FALSE, FALSE, 90.0, 3, 6),

(5, TRUE, TRUE, TRUE, 95.0, 0, 2);

#### **Insert Performance Metrics**

INSERT INTO performance\_metrics (employee\_id, days\_absent, productivity\_score, job\_satisfaction\_score)

VALUES

(1, 2, 89.7, 8.5),

(2, 0, 92.3, 9.0),

(3, 5, 85.6, 7.2),

(4, 3, 88.1, 7.8),

(5, 1, 93.5, 9.1);

### **3. Generate the Data with Queries**

#### **Example: Retrieve All Data with Relationships**

SELECT

e.employee\_id,

e.name,

e.sex,

e.date\_of\_birth,

e.job\_title,

h.health\_screening\_done,

h.safety\_training\_done,

h.ergonomic\_assessment\_done,

h.annual\_health\_checkup\_score,

h.workplace\_injury\_count,

h.stress\_level,

p.days\_absent,

p.productivity\_score,

p.job\_satisfaction\_score

FROM

employees e

JOIN

health\_safety\_metrics h ON e.employee\_id = h.employee\_id

JOIN

performance\_metrics p ON e.employee\_id = p.employee\_id;

### **4. Analyzing the Data**

**To generate the summary statistics and insights:**

#### **Average Productivity and Satisfaction by Job Title**

SELECT

e.job\_title,

AVG(p.productivity\_score) AS avg\_productivity\_score,

AVG(p.job\_satisfaction\_score) AS avg\_job\_satisfaction\_score

FROM

employees e

JOIN

performance\_metrics p ON e.employee\_id = p.employee\_id

GROUP BY

e.job\_title;

#### **Correlation Between Stress Level and Productivity**

SELECT

h.stress\_level,

AVG(p.productivity\_score) AS avg\_productivity\_score

FROM

health\_safety\_metrics h

JOIN

performance\_metrics p ON h.employee\_id = p.employee\_id

GROUP BY

H.stress\_level;

### **Part 5: Integration and Testing**

#### **Integration: Documenting the Process of Importing Data into Excel and Ensuring Consistency**

**1. Preparing the Data:**

* **Step 1: Review the Dataset:**
  + Before importing the data into Excel, ensure that the dataset is clean and properly formatted. This includes checking for missing values, ensuring that all data types are consistent (e.g., dates are in date format, numbers in numeric format), and verifying that there are no duplicated entries.
* **Step 2: Save the Dataset in a Compatible Format:**
  + Save the dataset as a CSV file or another format compatible with Excel. This ensures easy import and reduces the chances of compatibility issues.

**2. Importing Data into Excel:**

* **Step 1: Open Excel and Access the Import Wizard:**
  + Launch Microsoft Excel, then navigate to the “Data” tab on the ribbon.
  + Click on “Get Data” and select “From Text/CSV” if importing a CSV file.
  + Browse and select the dataset file, then click “Import.”
* **Step 2: Preview and Adjust Import Settings:**
  + Excel will display a preview of the dataset, showing how the data will be imported.
  + Check the delimiter (e.g., comma, semicolon) and adjust it if necessary.
  + If the preview looks correct, click “Load” to import the data into a new worksheet.

**3. Ensuring Data Consistency:**

* **Step 1: Validate Column Headers:**
  + Verify that column headers are correctly labeled and match the intended data fields.
  + Make adjustments if necessary to ensure clarity and consistency across the dataset.
* **Step 2: Data Type Verification:**
  + Ensure that each column is formatted correctly (e.g., dates are in date format, numbers are in number format).
  + Utilize Excel’s “Format Cells” option to adjust any misformatted data.
* **Step 3: Handle Missing Data:**
  + Identify any missing values and address them appropriately. This may involve filling in missing data, using Excel’s functions like IFERROR, or noting where data is unavailable.
* **Step 4: Standardize Data Values:**
  + Check for and standardize variations in data entries (e.g., "Yes" vs. "Y", "New York" vs. "NYC").
  + Use Excel’s “Find and Replace” tool to correct any inconsistencies.
* **Step 5: Cross-Check Against Source Data:**
  + Cross-check a sample of the imported data against the original dataset to ensure accuracy.
  + Use Excel functions like VLOOKUP to verify that key data points have been imported correctly.

#### **Testing: Test the Integration and Functionality of Your Excel Dashboard**

**1. Initial Testing of Data Integrity:**

* **Step 1: Create Sample Pivot Tables:**
  + Generate pivot tables to summarize key data points. This will help identify any issues with data aggregation and ensure that all data fields are functioning correctly.
* **Step 2: Verify Calculations:**
  + Test Excel formulas and calculations (e.g., sum, average, ratios) on the dataset to ensure they return the expected results.
  + Use Excel’s “Evaluate Formula” tool to trace and verify more complex calculations.

**2. Dashboard Functionality Testing:**

* **Step 1: Interactive Elements:**
  + Test any interactive elements within the dashboard (e.g., slicers, drop-down menus) to ensure they function correctly.
  + Ensure that changing one element updates all relevant charts, graphs, and tables accordingly.
* **Step 2: Data Filtering and Sorting:**
  + Apply various filters and sorting options to test how the data is displayed and ensure the dashboard responds correctly to different user inputs.
  + Ensure that filtered data maintains consistency across all visualizations.
* **Step 3: Visual Consistency and Clarity:**
  + Review all visual elements (e.g., charts, graphs) for consistency in color schemes, labeling, and data representation.
  + Ensure that all visualizations are clear, easy to understand, and accurately represent the underlying data.

**3. Final Verification:**

* **Step 1: Cross-Check Insights:**
  + Verify that the insights generated by the dashboard align with your initial expectations and the goals of the analysis.
  + Compare the dashboard outputs with manual calculations to ensure accuracy.
* **Step 2: User Testing:**
  + If possible, have a colleague or another stakeholder review the dashboard to identify any usability issues or areas for improvement.
* **Step 3: Documentation:**
  + Document any issues found during testing and describe how they were resolved.
  + Provide guidance on how to use the dashboard effectively, including tips for interacting with the data.