**Data Analytics – Deloitee Virtual Internship**

**Daikibo Project Notes**

**Task 1: Data Analysis Using Tableau (Machine Telemetry)**

**Question:**

Task 1: Data analysis

Here is the background information on your task

Using a data unification algorithm, the tech team at our client, Daikibo, has converted all telemetry data collected from its 4 factories:

Daikibo Factory Meiyo (Tokyo, Japan)

Daikibo Factory Seiko (Osaka, Japan)

Daikibo Berlin (Berlin, Germany)

Daikibo Shenzhen (Shenzhen, China)

Each location has 9 types of machines, sending a message every 10 mins. Daikibo has been collecting this data for one month (May 2021) and they've shared this data in the form of a single JSON file.

The reason the client wanted to collect telemetry was to answer 2 questions:

In which location did machines break the most?

What are the machines that broke most often in that location?

Here is your task

Your task is to analyse the telemetry data collected by Daikibo in a software called Tableau. Here are the steps that you need to take:

Download the free trial of Tableau (link in the Resources).

Install Tableau on your computer and register an account with the same email you used to download the software.

Download the daikibo-telemetry-data.json.zip file -> unzip -> and import it in Tableau.

Create a calculated measure field called "Unhealthy" with a value of 10 for every unhealthy status (representing 10 mins of potential down time since the previous message).

Create a bar chart called “Down Time per Factory”.

Create a new sheet with a new bar chart called “Down Time per Device Type”.

Create a Dashboard with the 2 previous sheets and set the first chart to be used as a filter (selecting a factory in the first chart shows only the down time of the machines in this factory in the second chart).

Select the factory with the most down time (click on its bar), make a screenshot of the dashboard and upload it as a submission for this task.

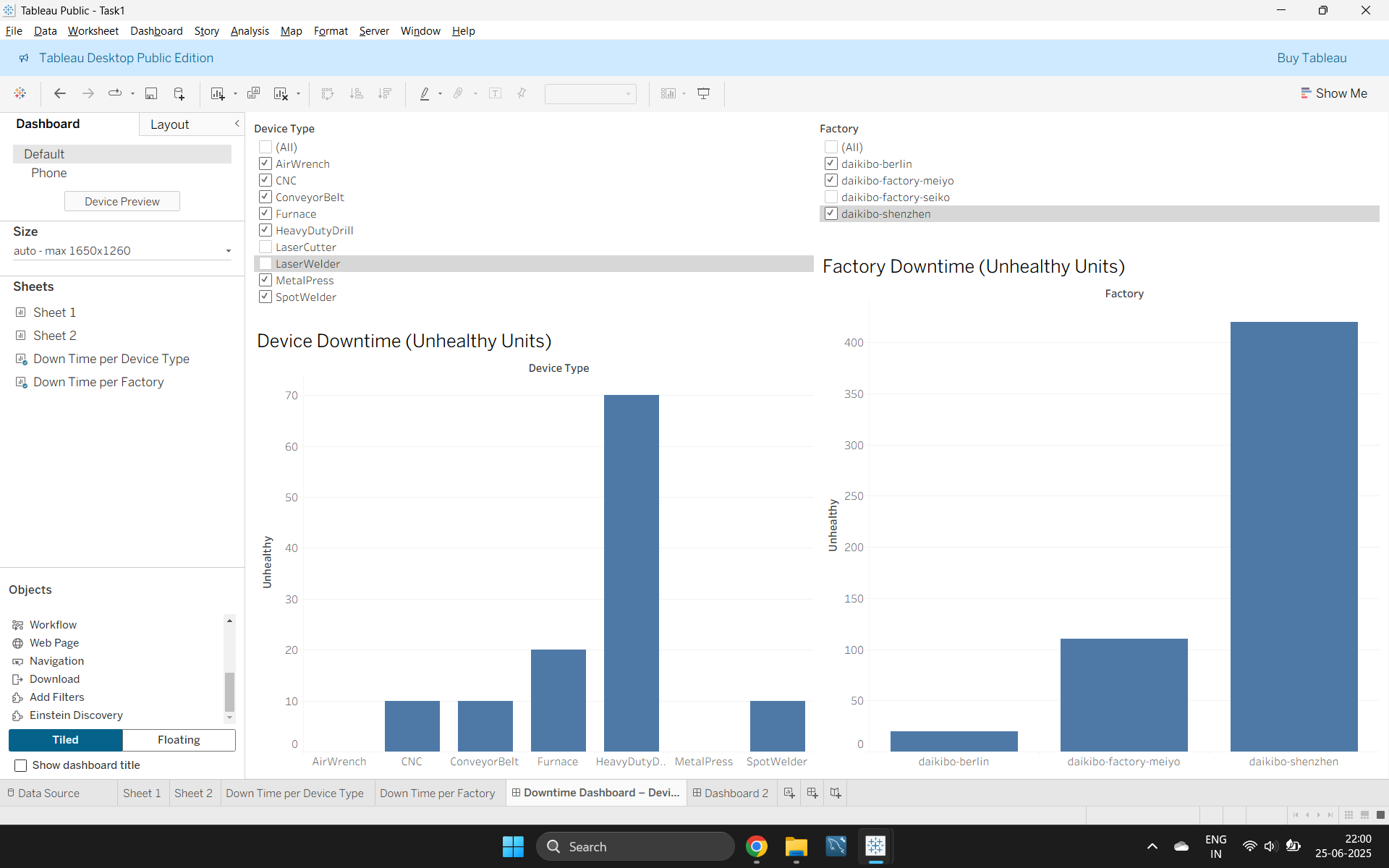
Please find a step-by-step Task Guide pdf attached in the resources below to help you out.

**Background:**

* Telemetry data collected from 4 factories for **May 2021**:
  + Meiyo (Tokyo, Japan)
  + Seiko (Osaka, Japan)
  + Berlin (Germany)
  + Shenzhen (China)
* 9 machine types per factory → data sent every 10 minutes
* Goal: Identify the **location with the most machine breakdowns** and the **specific machines** responsible.

**Step-by-Step Guide:**

1. **Download & Set Up:**
   * Install **Tableau Public** or free trial version. (https://www.tableau.com/products/desktop/download)
   * Sign in using the same email used for download.
2. **Load the Data:**
   * Download and unzip daikibo-telemetry-data.json.zip.
   * Open Tableau → Select JSON as the data source → Import the file.
3. **Create a Calculated Field (Unhealthy):**
   * Go to **Data Pane → Right-click → Create Calculated Field**.
   * Name: Unhealthy
   * Formula:
   * IF [status] = 'unhealthy' THEN 10 ELSE 0 END
   * Reason: Each unhealthy status = 10 minutes downtime
4. **Create Sheet 1: “Down Time per Factory”**
   * Rows: Factory
   * Columns: SUM(Unhealthy)
   * Mark Type: Bar Chart
5. **Create Sheet 2: “Down Time per Device Type”**
   * Rows: Device Type
   * Columns: SUM(Unhealthy)
   * Mark Type: Bar Chart
6. **Create Dashboard:**
   * Add both sheets to a new dashboard.
   * Make Sheet 1 (Factory) **a filter**:
     + Click on the sheet → click on filter icon → “Use as Filter”
7. **Interact and Identify:**
   * Click the bar for the **factory with the highest downtime**.
   * The second chart now shows **top failing machines** for that factory.
8. **Screenshot & Submit:**
   * Take a screenshot of the dashboard.
   * Upload it as your submission for Task 1.



**Task 2**

Here is the background information on your task

After a worrisome number of internal complaints about gender inequality in terms of salary, Daikibo Industrials wants us to help them investigate.

The Forensic Tech team has built an algorithm to quantify “level of gender pay equality” for most job roles within the company, in all company locations. Our Forensics lead thinks it would be a great idea for you to finish the job.

Task 2: Forensic technology

Here is your task

We have processed all data on employee compensation and generated an Excel file (Equality Table.xlsx, available in the Resources) containing 3 columns:

Factory

Job Role

Equality Score (integer; ranging between -100 and +100; 0 is ideal)

Here is your task:

Create a 4th column (Equality class), classifying the equality score into 3 types:

Fair (+-10)

Unfair (<-10 AND >10)

Highly Discriminative (<-20 AND >20)

Examples:

10 → Fair

-9 → Unfair

-30 → Highly Discriminative

Please find the Equality Table you need to edit in the resources below. When you are done, upload the edited version of the file.

**Task 2: Gender Pay Equality Classification (Excel)**

**Background:**

* Data source: Equality Table.xlsx
* Contains:
  + Factory
  + Job Role
  + Equality Score (range: -100 to +100; 0 = ideal equality)

**Step-by-Step Guide:**

1. **Download the file Equality Table.xlsx**
2. **Open it in Excel or Google Sheets**
3. **Create a New Column:**
   * Add a 4th column called Equality Class
4. **Apply Classification Logic:**  
   Use this rule to fill the column:
   * **Fair** → Score between -10 and +10 (inclusive)
   * **Unfair** → Score between -20 and -11 OR +11 and +20
   * **Highly Discriminative** → Score < -20 OR > +20
5. **Use Excel Formula:**  
   Assuming Equality Score is in Column C (starting from row 2):
6. =IF(ABS(C2)<=10, "Fair", IF(ABS(C2)<=20, "Unfair", "Highly Discriminative"))
7. **Drag to apply formula to all rows**
8. **Double-check edge cases**
9. **Save and Submit** the updated Excel file.

**Tools Used:**

* **Tableau** (for interactive dashboard analysis)
* **Excel** (for forensic classification using logic)