About the Project

The project seeks to do descriptive data analysis on employee data. The the results of the analysis could then be used to inform diagnostic analysis to explain trends and observations and prescriptive analysis

About the Data

Name: Computer Login
Source: Udemy Course
Fields:

1. User name - Text
2. Log in - [Date-Time]
3. Log out - [Date-Time]
4. Time zone - [Text] - Shows difference from GMT
5. Role - [Text] - Indicates the capacity at which the employee serves the company

Generated Data

1. Duration for every log in
2. Work duration per employee per day
3. The period of day within which every log in occurs
4. The number of logged-on users at any given hours
5. Average working hours per employee per day
6. Average working hours per week

What insights can we get from the data?

Employees

1. How many employees are in the company?
2. How many employees are in each department?

Work hours & Overtime

1. Average work hours per employee/role
2. Overtime patterns/trends
3. Employees frequently working overtime
4. Overtime frequency and amount across departments

Facility Utilization

1. Peak Workload Hours

Shift Patterns

1. Distribution of employees across different shifts

Work-Life Balance

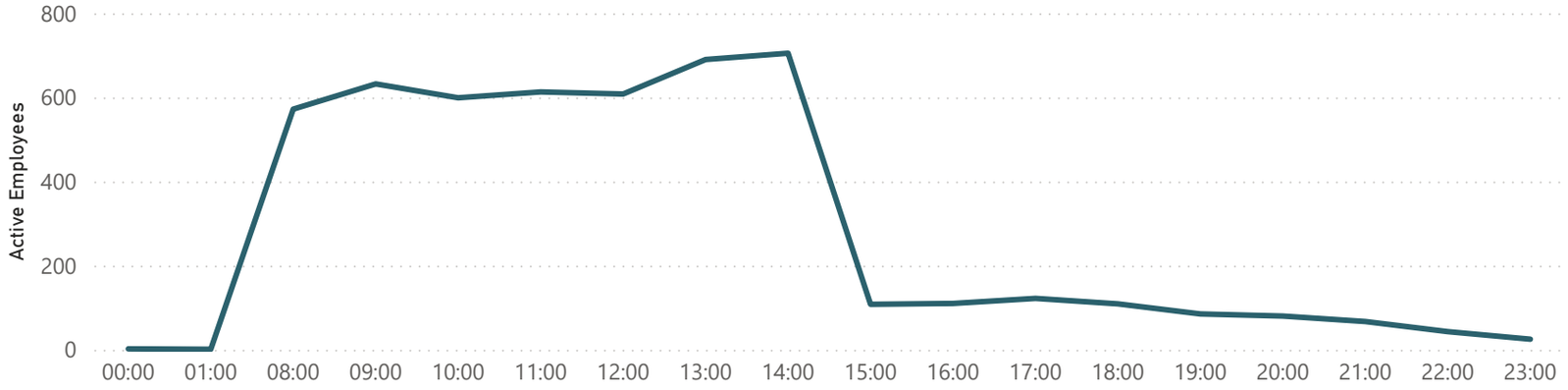
1. Average work hours per week
2. Employees with excessive work hours

Summary of Employees

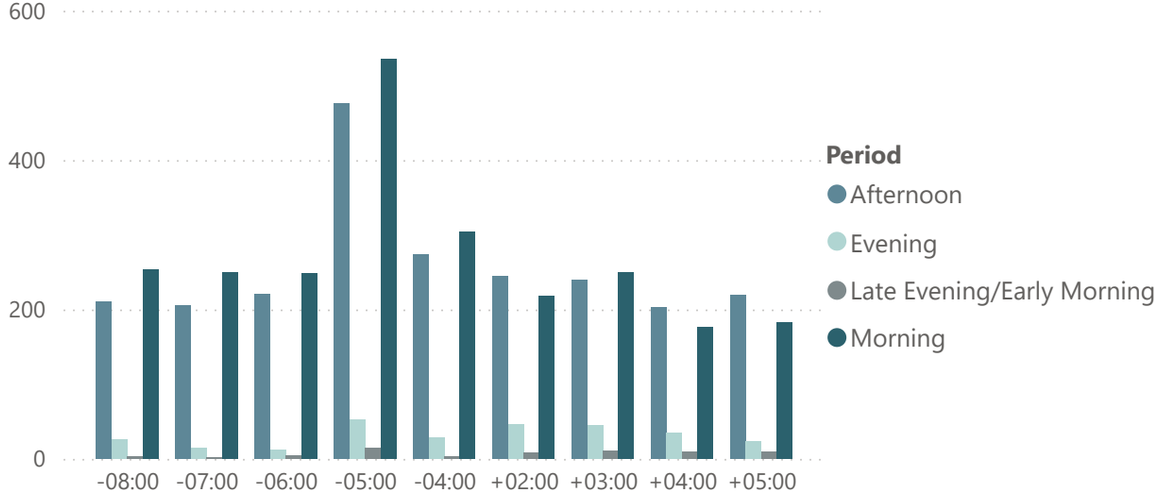
The company is largely comprised of Data Analysts. Most employees regardless of time zone, tend to work during the **morning and afternoon shifts**. The **peak work load period** occurs between **08:00** and **14:00**. During these times, System performance (e.g. servers and WiFi) would require to be optimal or boosted.

Role
- Data Analyst
- Supervisor

System Load

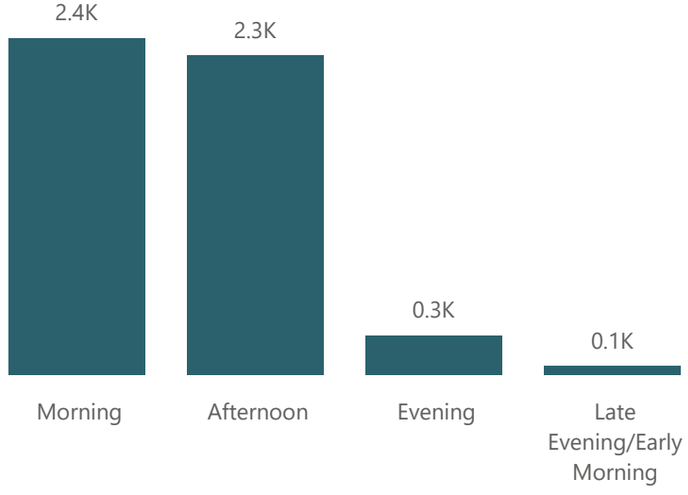


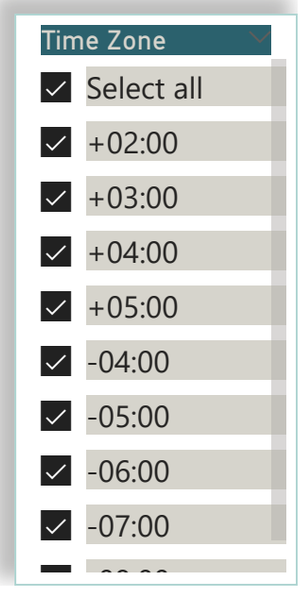
Log-In Distribution according to Timezones



Period
- Afternoon
- Evening
- Late Evening/Early Morning
- Morning

Log In Activity Across Shifts





Summary of Work Hours

The **average work hours** is well **below 8 hours**. This could be an indication of **low productivity**. **Sunday**, has the highest average number of working hours.

User Name	Average Weekly Hours
Harper	11.59
Charlotte Martin	11.45
Abigail Jane Bailey	11.20
Evelyn Butler	11.07
Sophia Ramirez	10.97
Isabella Adams	10.89
Emma Harris	10.83
Elizabeth Jenkins	10.75
Ava Jenkins	10.50
Emily Nelson	10.14
Olivia Sanchez	9.98
Claire Brooks	9.98
Amelia D. Rivera	9.73
Savannah	9.59
Penelope Powell	9.58
Ella Clark	9.38
Mia Wilson	9.27
Avery Barnes	9.13
Zoey Roberts	9.02
Luna Cox	8.98
Violet Stewart	8.81

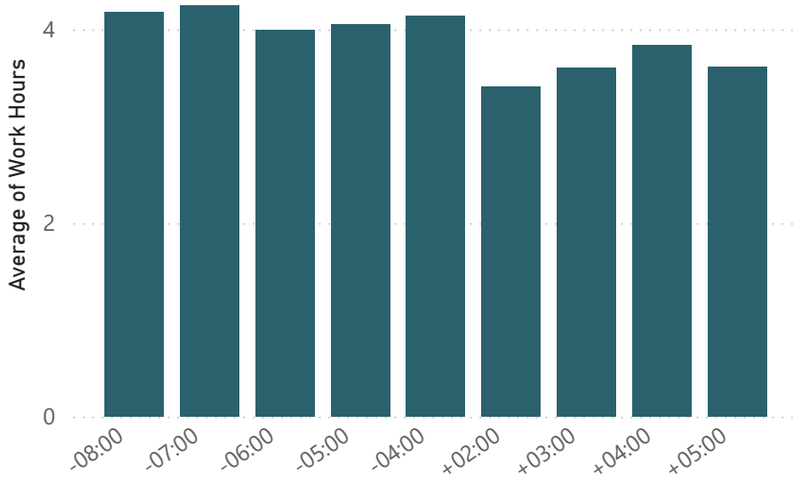
4.87
Average Work Hours per Week

4.85
Average Work Hours Per Day

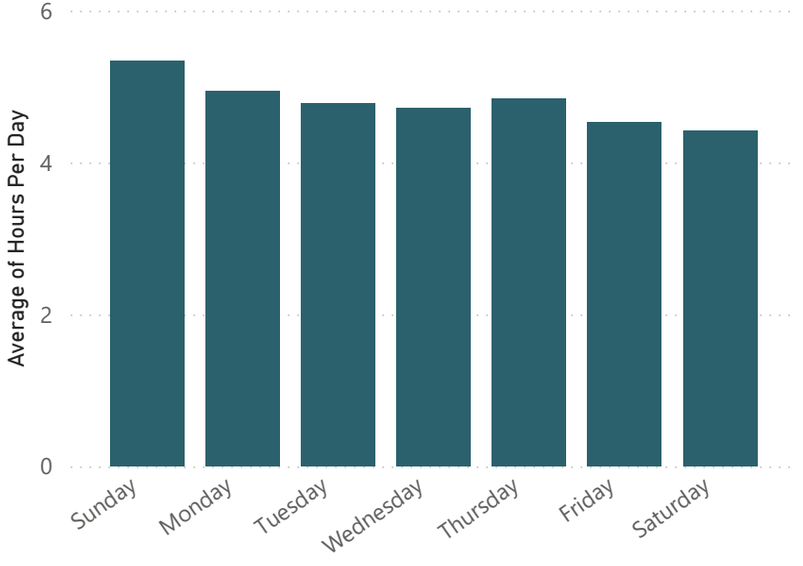
Employee
All

Role
All

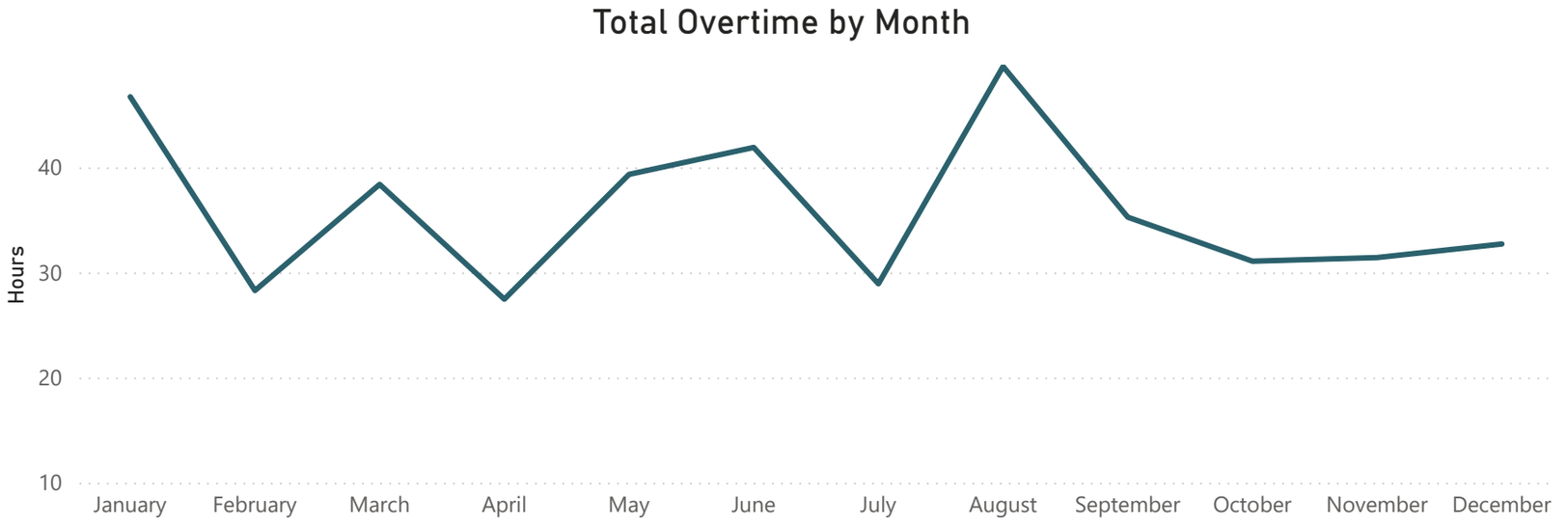
Average Work Hours per Time zone



Average Work Hours Per Day

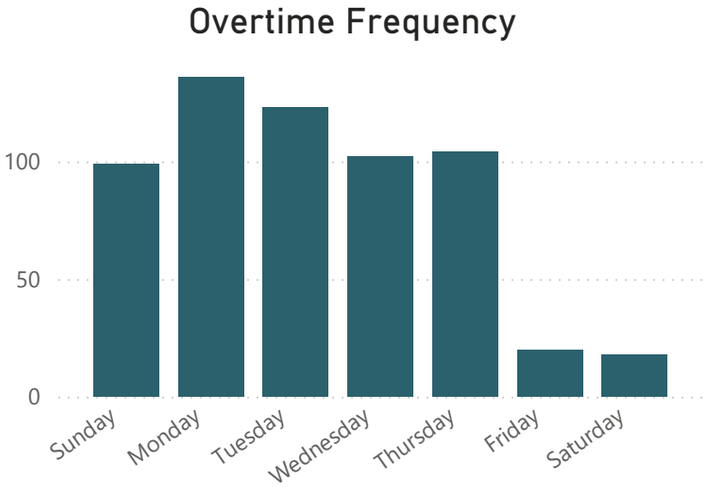


Overtime Hours



431

Total Overtime Hours

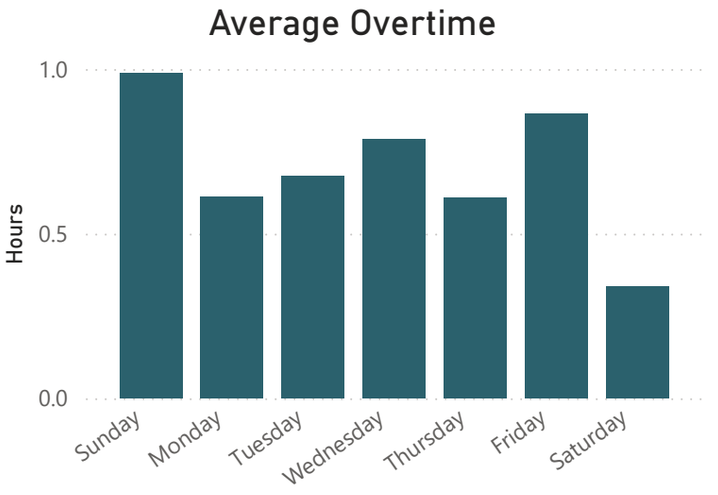


Employees with Highest Overtime Frequency

User Name	Overtime Occurrences
Ava Jenkins	23
Charlotte Martin	20
Isabella Adams	20
Abigail Jane Bailey	19
Elizabeth Jenkins	17
Emily Nelson	16
Emma Harris	16
Harper	16
Savannah	16
Amelia D. Rivera	15

Employees with Highest Total Overtime Hours

User Name	Total Overtime
Abigail Jane Bailey	26.56
Charlotte Martin	20.24
Elizabeth Jenkins	19.93
Ava Jenkins	19.15
Isabella Adams	16.63
Savannah	15.90
Emma Harris	14.95
Emily Nelson	14.46
Harper	10.62
Luna Cox	10.38



Concerns with the Data

- **Log In and Log Out date-times:** For some data points the log out occurs earlier than the log in. This is practically impossible since the logs contain both date and time stamps. Some log-on periods are also very short, while this is practically possible, the previous statement already raises concerns on accuracy of data.
- While the data indicates that employees work in different time zones, assuming the data is collected automatically from digital devices, and that for any particular log-on, the date-time stamp is recorded from the same device, the chances of such an error occurring should be minimal.
- This may mean the work durations recorded may largely be inaccurate hence affecting results of the analysis. For example, the average work hours are quite low at about 5 hours.

Data that would be other wise useful but not included

Note: This considers the fact that this project was as a personal initiative to further explore data that was provided for completion of a much simpler exercise (i.e. transforming and adding columns in Power BI). As such, concerns raised and recommendations given are based on the assumption that this was the original data provided

The following data would help get even more meaningful analysis:

1. More on employee details: Age, gender, Country of residence
2. Projects undertaken and periods within which they occurred

Questions for Diagnosis

- Why does Sunday have highest average overtime and work hours?
- Are employees working this job as their part-time or full-time?
- Is there a specific department often working overtime?
- Overtime hours seem to always peak and fall. Do these indicate the start and end of major projects?
- Is there a different remuneration system for overtime hours? Are there serious implications on the amount of overtime incurred? For example in the case of employees paid by the hour or higher pay for over time hours.
- What are the reasons for overtime? Could these reasons be recorded?
- Is the amount of hours worked a good indication of productivity or could we come up with better metrics, i.e. milestones achieved?
- Why is the morning to afternoon shift more popular?