

# Call Center Analysis

By: Nadav Erez

Data Mining course - Dr. Chen Hajaj

[Link to code - Google Colab \(viewer\)](#)



אוניברסיטת  
אריאל  
בשומרון

# Roadmap

- ❖ Introduction
- ❖ Correlation
- ❖ Target variables shown over time
- ❖ Top performing teams and agents
- ❖ Two-way ANOVA
- ❖ Summary



© marketoonist.com

# Introduction

The dataset that this analysis is based on contains data collected from 3 different sites of a call center.

The dataset's main attributes:

Factors / Independent Variables:

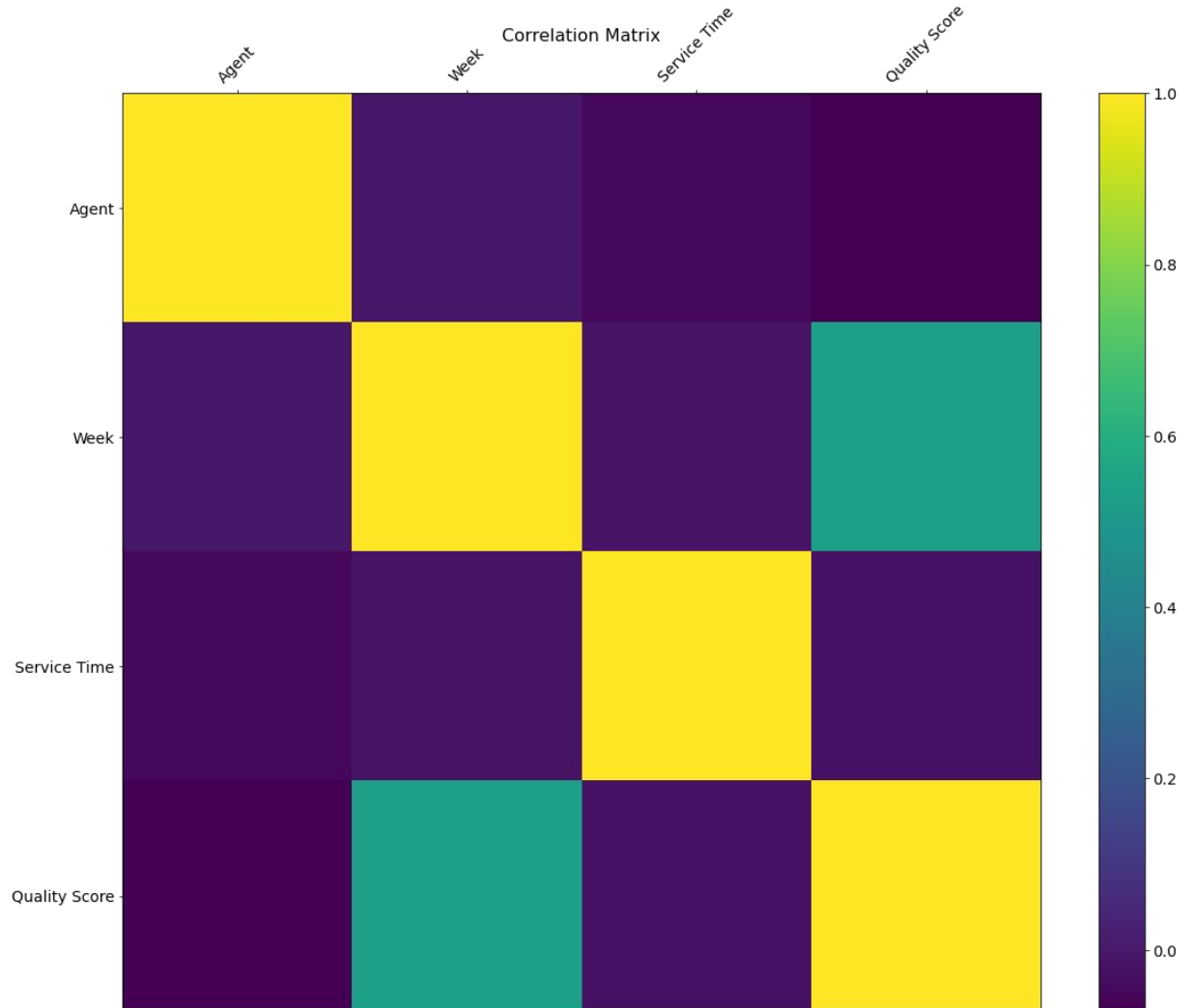
- **Site** - East, North & South
- **Client** - A & B
- **Supervisor** - Team leader
- **Agent** - Employee
- **Week** - 1-5

Dependent / Target Variables:

- **Service Time** - float, seconds
- **Quality Score** - float, range 1-10

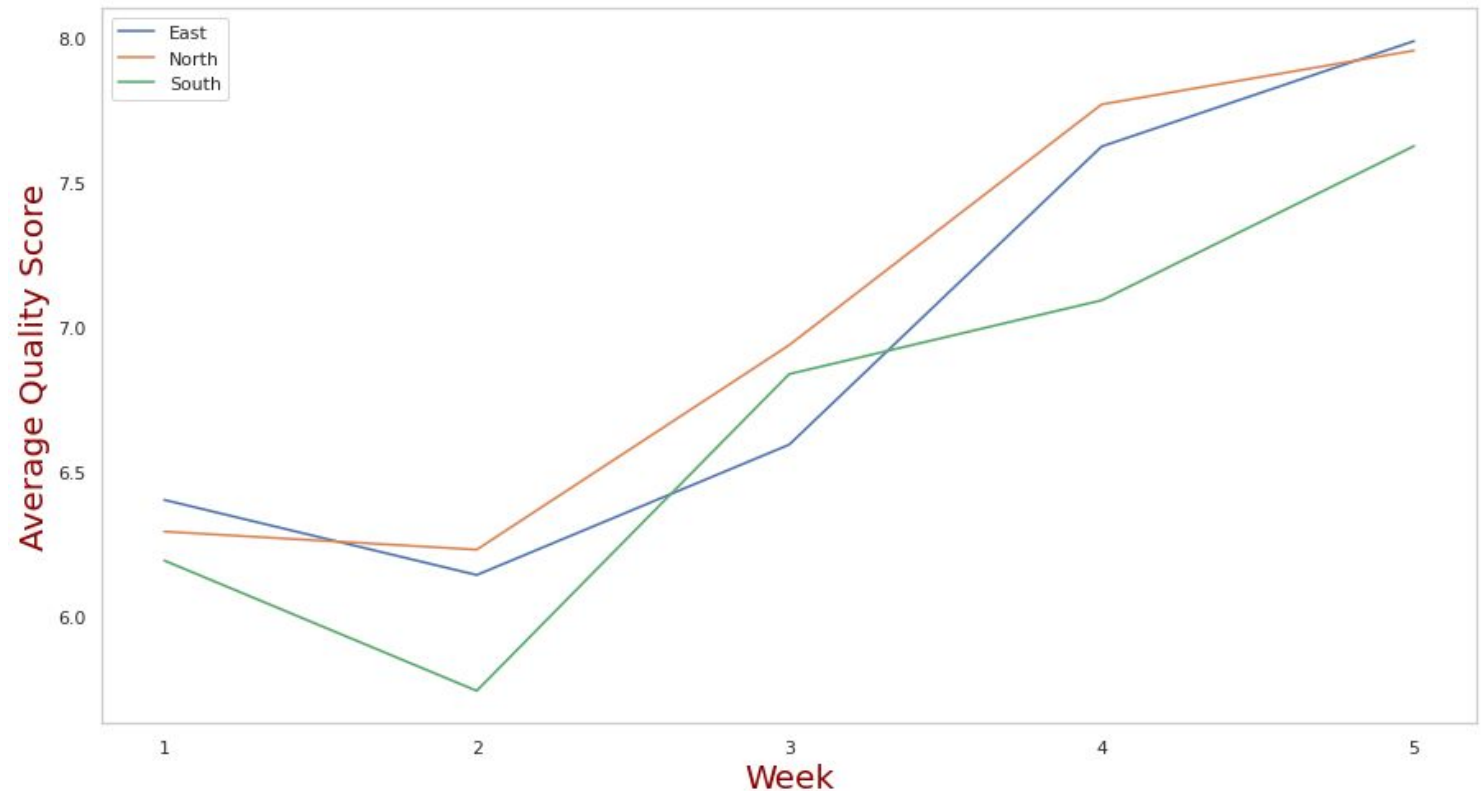
# Correlation

Overall the data has low correlation scores, indicating that there is no linear correlation between the different variables. However, highest correlation score is between the Quality Score and the week. This aligns with our finding that the average quality score grows over time across the whole call center.



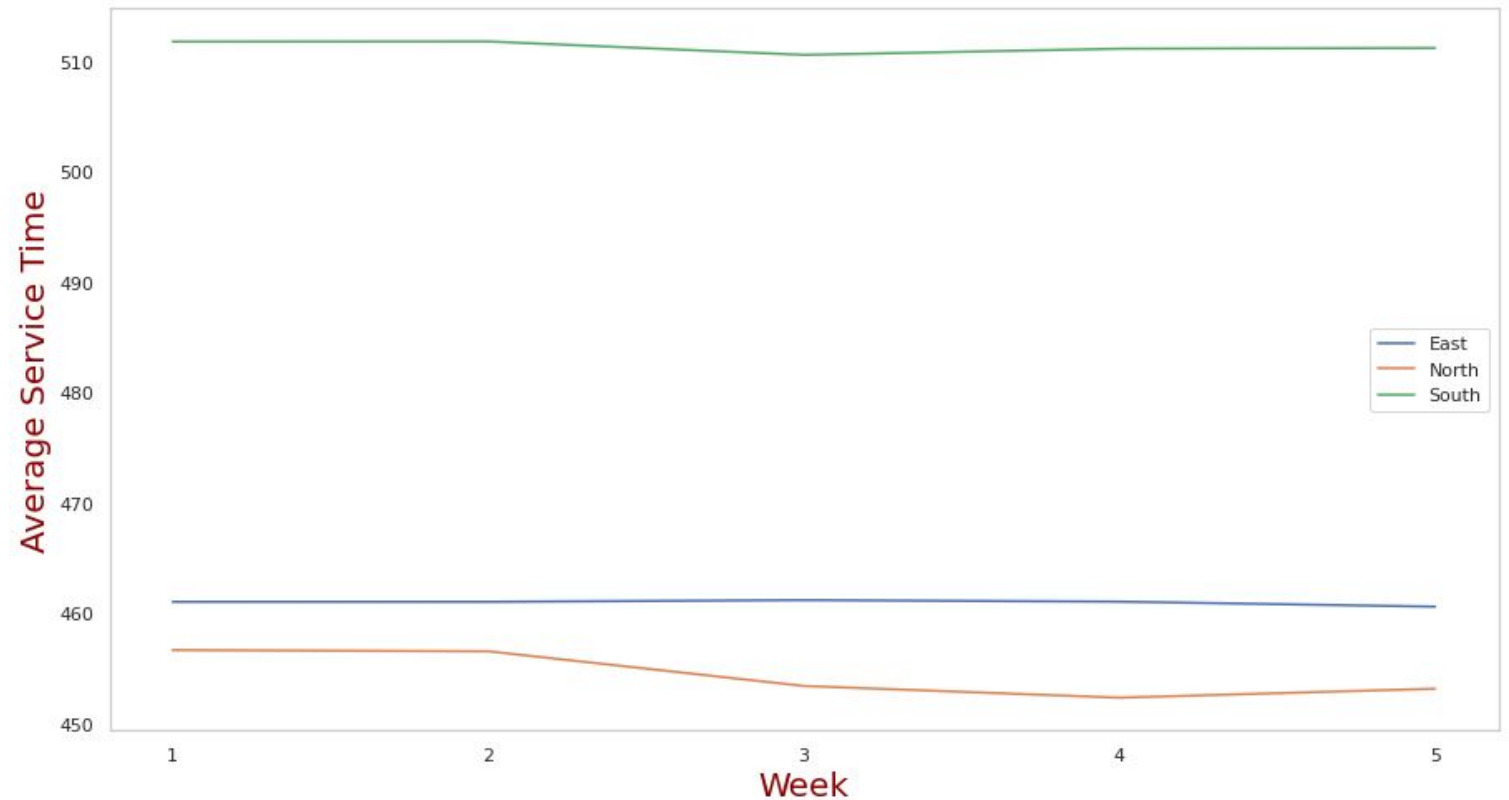
# Quality Score by Site

- Average Quality Score gradually increases over the time period the data was collected in.
- In terms of quality, looks like the North site outperforms the rest.

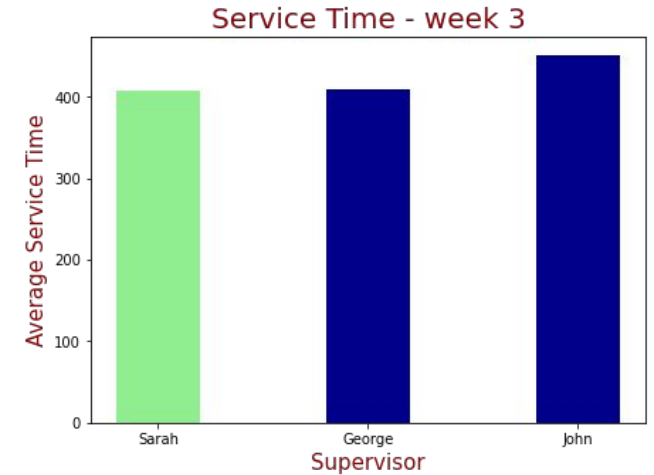
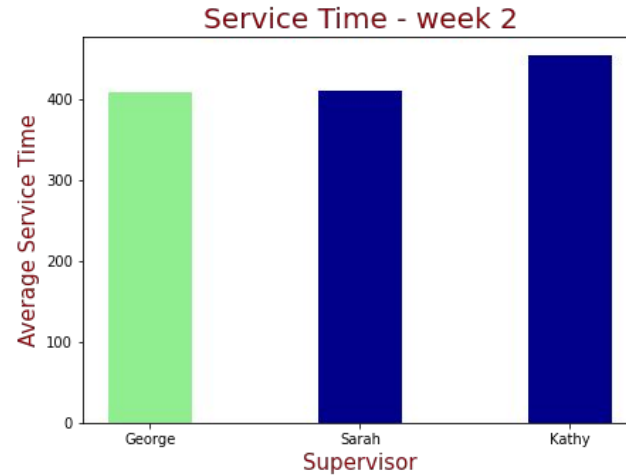
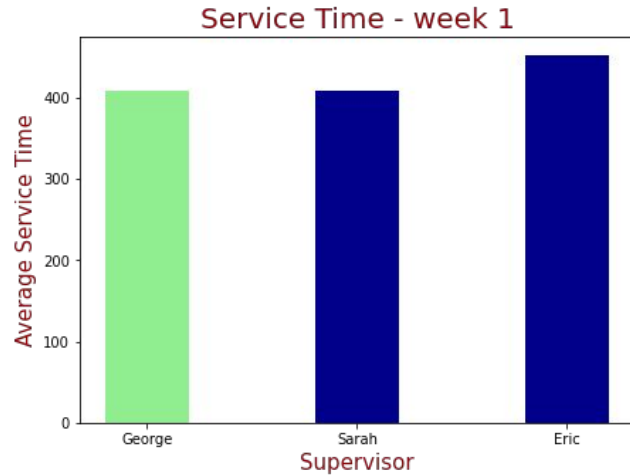


# Service Time by Site

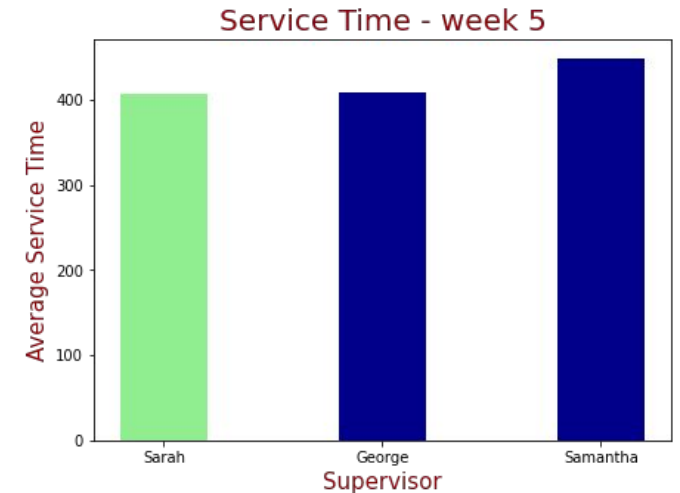
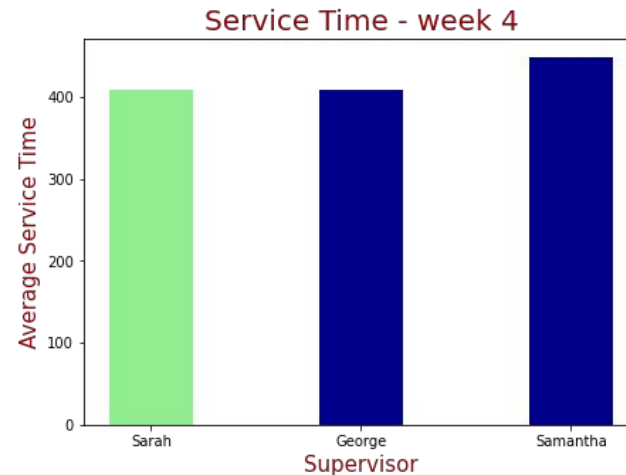
- Average Service Time stays relatively the same over the time period the data was collected in.
- In this case as well - the North site wins with the lowest Service Time recorded.



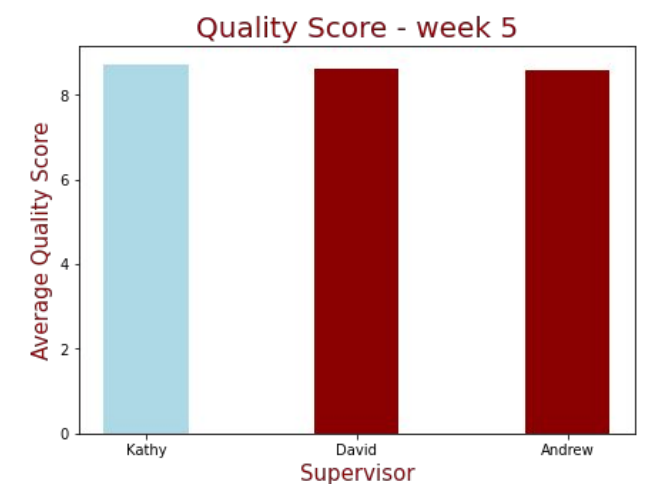
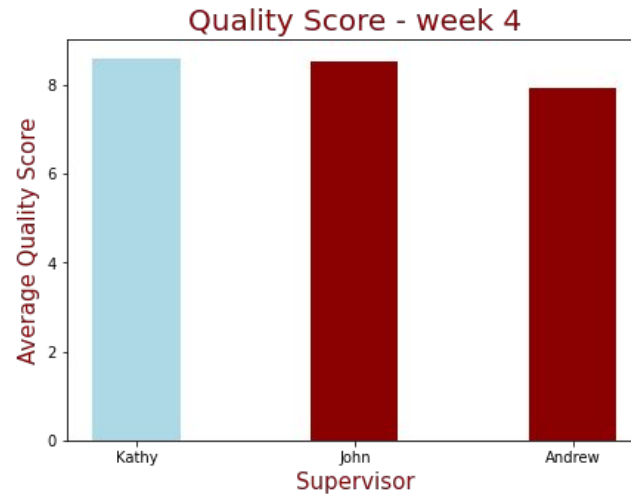
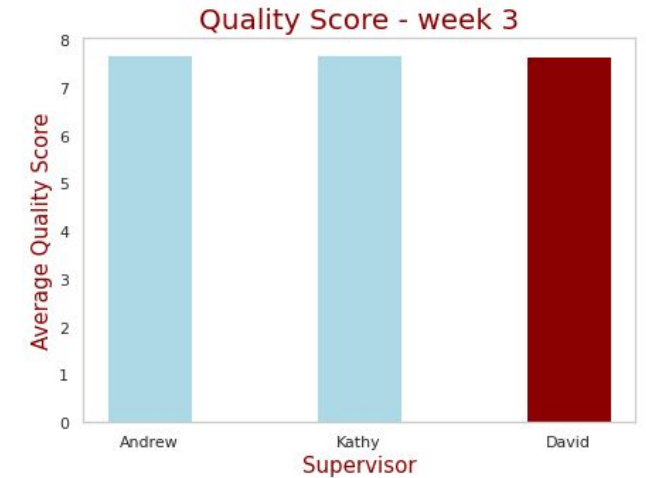
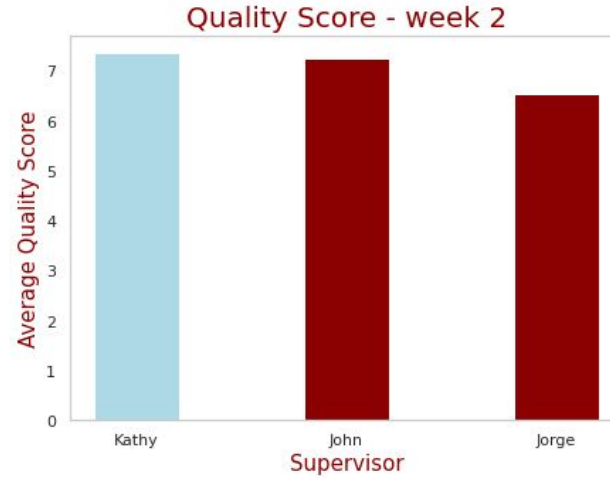
# Top Performing Teams - Lowest Service Time



George and Sarah's teams perform best with lowest average Service Time in the whole period



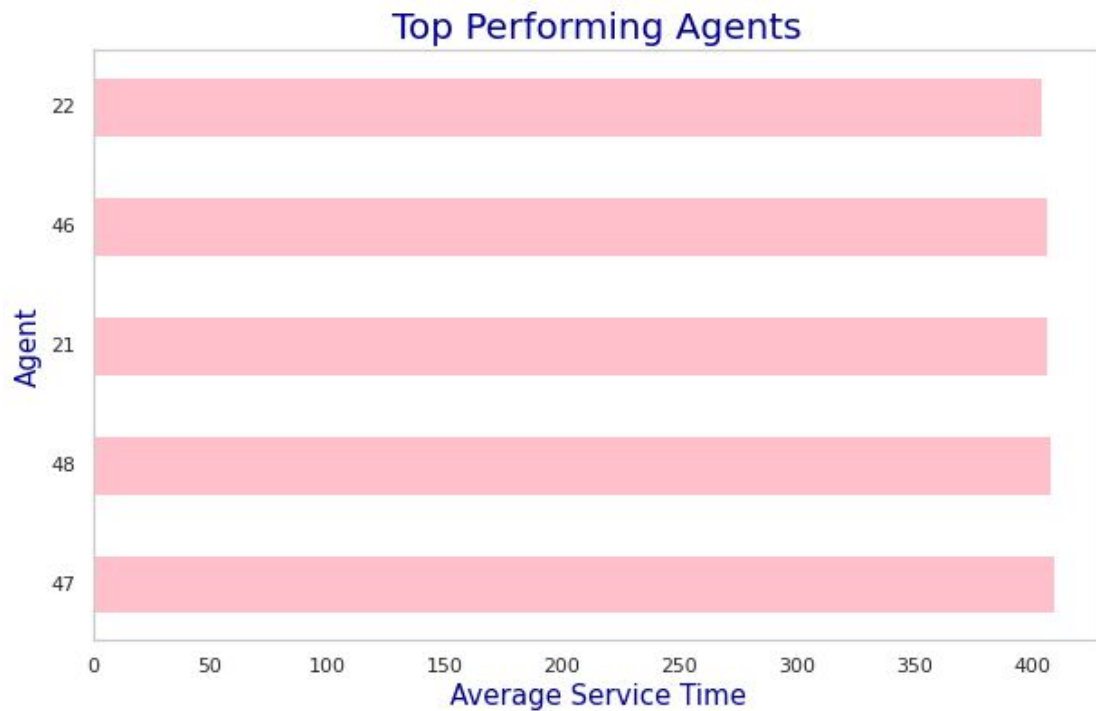
# Top Performing Teams -Highest Quality Score



David, Kathy and Andrew's teams perform best with highest average Quality Score



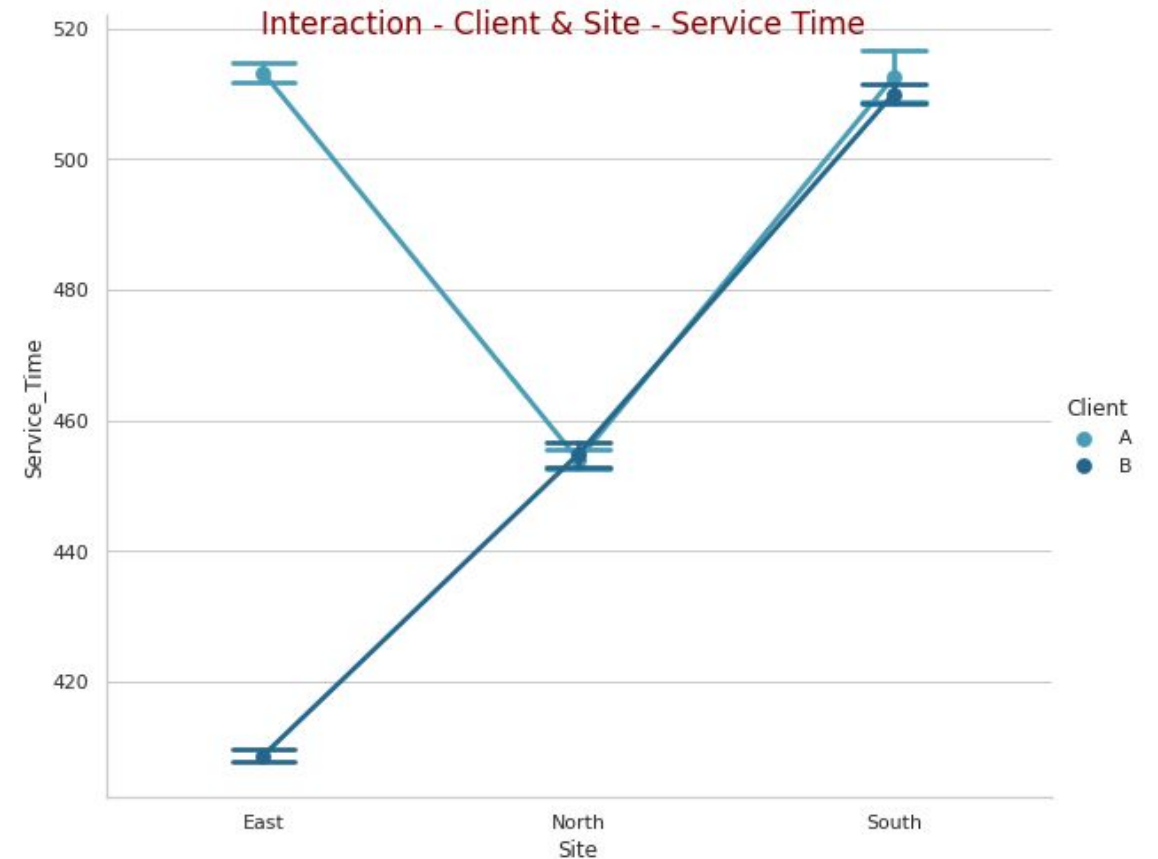
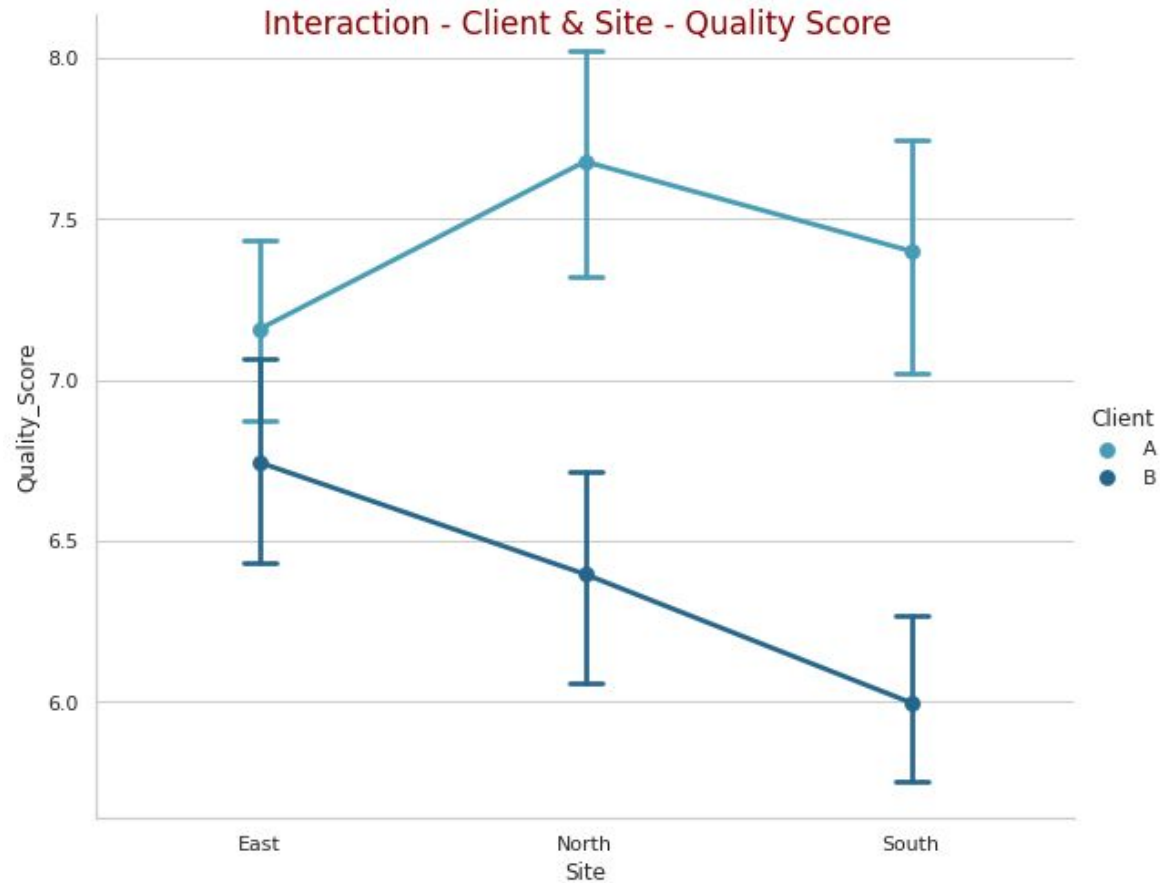
# Top 5 Employees Overall



Best employees - agents 22 and 25

# Analysis of Variance

In this part, a 2-way ANOVA is performed to test significant differences in Service Time and Quality Score between the different sites and clients.



# Hypotheses

## **Two-way ANOVA Hypotheses:**

Quality Score / Service Time

H0: There is no difference in the mean quality score / service time between the different sites

H1: At least one mean is in fact different

H0: There is no difference in the mean quality score / service time between both clients A & B

H1: The mean is different between the two clients

H0: The interaction between site and client has no effect

H1: The interaction has an effect

# Results

## Quality Score

index	sum_sq	df	F	PR(&gt;F)
Site	4.953062500000205	2.0	2.150741455277539	0.11869328890740799
Client	63.98919010416651	1.0	55.5713576586851	1.74075325612084e-12
Site:Client	11.663520833333147	2.0	5.064587368066078	0.007026716041174749
Residual	269.44582812499993	234.0	NaN	NaN

From the analysis we can see that there is a significant difference between clients regarding quality score, and that the interaction does in fact have an effect on quality time.

## Service Time

index	sum_sq	df	F	PR(&gt;F)
Site	155033.25624999977	2.0	1605.4031652283709	2.2385792801781353e-137
Client	75056.90859375014	1.0	1554.461301314029	2.54208259762569e-105
Site:Client	143468.72500000008	2.0	1485.6499231033877	1.0271214066182645e-133
Residual	11298.651562499997	234.0	NaN	NaN

Regarding service time, we reject all the null hypotheses and can say that there is a significant difference in the mean service time between the sites and clients, and that their interaction has an effect.

*That's all Folks!*