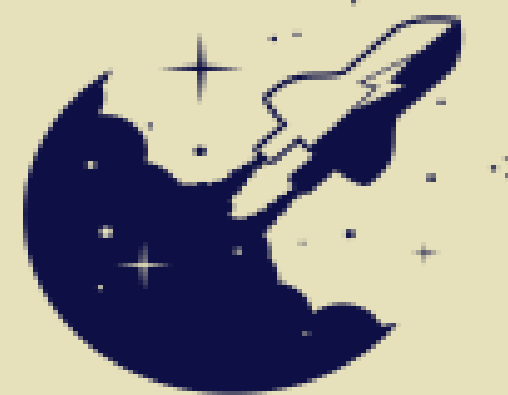


ARIEL  
UNIVERSITY

UNIVERSITY  
ARIEL



Fictional Company LTD  
Data Mining

# Data Analysis

16.03.2022

מציג: ליעד בן יחיאל - 207637414







# תוכן עניינים





# הקדמה - רקע החברה

Fictional Company LTD הינה חברה פיקטיבית למתן שירות לא ידוע.

החברה עובדת ב3 אזורים: צפון, דרום ומזרח.

לחברה קיימים 2 לקוחות בלבד A,B שעל כל שירות בה אחראי סוכן שמעליו נמצא מפקח.

הדאטה נבנה על סמך 5 שבועות ומדדי הביצוע של החברה ככלל ושל הסוכנים

והמפקחים בפרט נמדדים על פי זמן מתן השירות ורמת איכות השירות.

דוגמא לנתונים:

	Site	Client	Supervisor	Agent	Week	Service_Time	Quality_Score
0	East	A	Brian	41	1	509.0	6.5
1	East	A	Brian	41	2	505.0	6.9
2	East	A	Brian	41	3	NaN	5.9
3	East	A	Brian	41	4	505.0	7.1
4	East	A	Brian	41	5	511.0	9.1





# הקדמה - מטרת המצגת

**החברה מעוניינת לזהות תבניות ודפוסי התנהגות  
בחברה לשם - שיפור ושימור, השבחת עובדים  
מצטיינים וטיפול בעובדים עם תפוקה נמוכה ואירועים**





# טיוב הנתונים - נתוני החברה

## דאטה החברה כולל 240 שורות

ניתן לראות שמתוך דאטה סך 240 שורות  
בעמודת זמן השירות חסרים 9 ערכים.  
ובעמודת דירוג האיכות חסרים 6 ערכים.

```
display(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 240 entries, 0 to 239  
Data columns (total 7 columns):  
#   Column                Non-Null Count  Dtype    
---  ---                  
0   Site                  240 non-null   object   
1   Client                 240 non-null   object   
2   Supervisor            240 non-null   object   
3   Agent                  240 non-null   int64    
4   Week                   240 non-null   int64    
5   Service_Time           231 non-null   float64   
6   Quality_Score          234 non-null   float64   
dtypes: float64(2), int64(2), object(3)  
memory usage: 13.2+ KB  
None
```

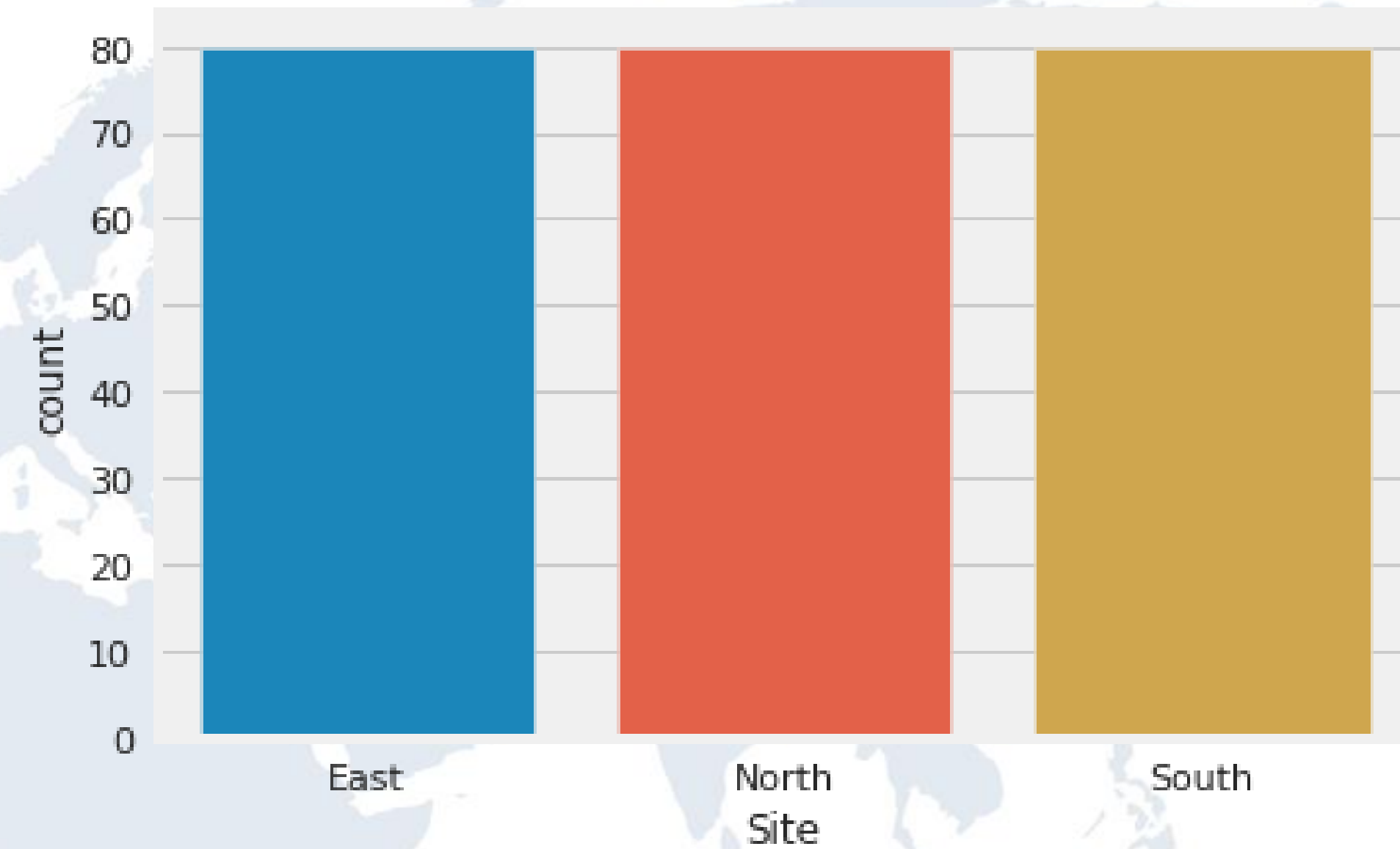
בקוד בוצע - הכנסת הערכים הממוצעים של  
העמודה היכן שהערכים חסרים.

```
mean_Service_Time 475.12554112554113  
mean_Quality_Score 6.895299145299149
```





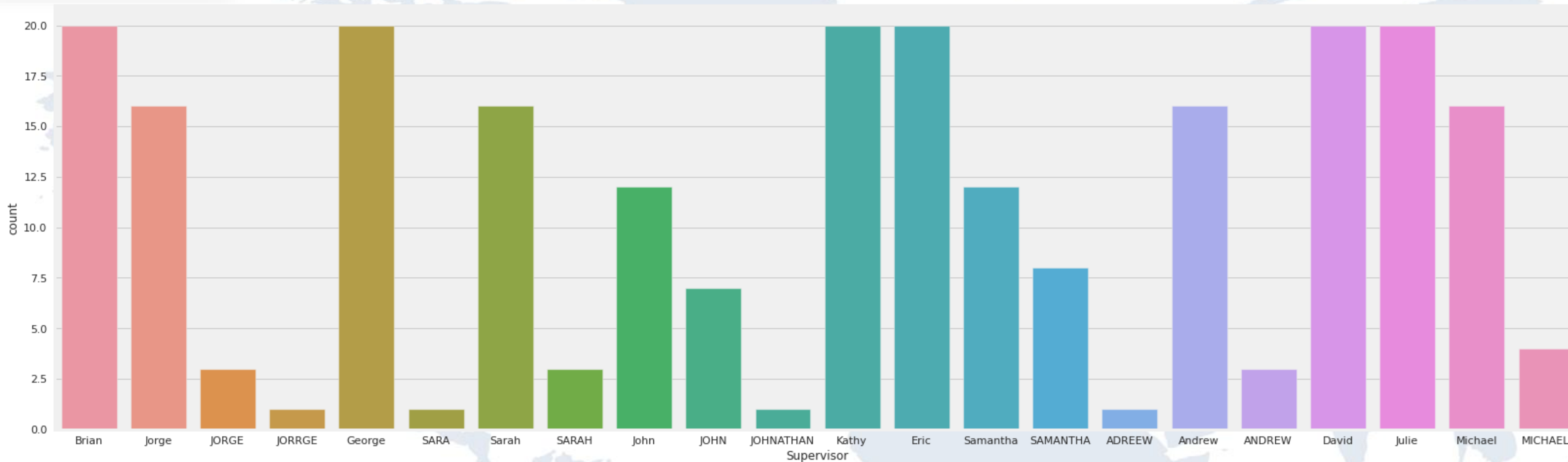
# טיוב הנתונים - בדיקת התפלגות הנתונים



גם בהתפלגות לפי אזור וגם כן בהתפלגות על פי לקוח ניתן לראות פילוג שווה וללא טעויות.



# טיוב הנתונים - בדיקת התפלגות הנתונים



ניתן לראות מספר טעויות הקלדה שנותנים רושם מוטעה של התפלגות הדאטה ולכן נחקור את התפלגות המפקחים לעומק.

## \*לשם תצוגה נכונה וצמצום טעויות הפכנו את כל השמות לאות ראשונה גדולה והשאר

### קטנות.

ניתן לראות בסקירה שישנן טעויות הקלדה שמטעות את המחשב לחשוב שישנם סוכנים חדשים.

ולכן כל שם שמופיע פעמיים או פחות ישתנה לשם המפקח בעל הדמיון הגדול ביותר אליו.

כפי שניתן לראות בקוד לצורך העניין (לא קריטי להבנה):

```
Counter({'adreew': 1,
        'andrew': 19,
        'brian': 20,
        'david': 20,
        'eric': 20,
        'george': 20,
        'john': 19,
        'johnathan': 1,
        'jorge': 19,
        'jorrge': 1,
        'julie': 20,
        'kathy': 20,
        'michael': 20,
        'samantha': 20,
        'sara': 1,
        'sarah': 19})
```

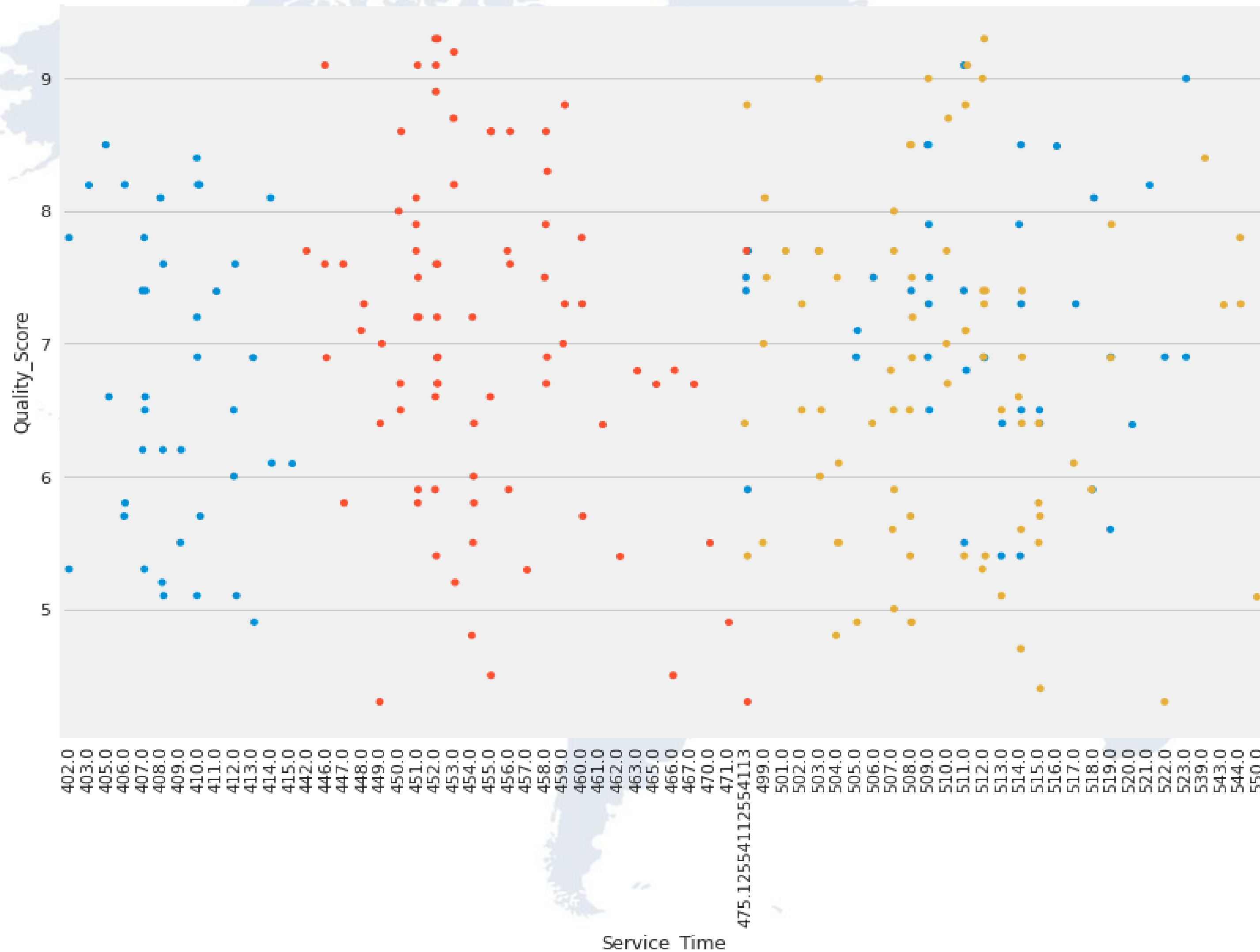
```
max_simily = 0
for word in df["Supervisor"]:
    max_simily_word = ""
    max_simily = 0
    for unique_word in listtt:
        simily = similar(word, unique_word)
        if simily >= max_simily:
            max_simily = simily
            max_simily_word = unique_word
    df["Supervisor"] = df["Supervisor"].replace(word, max_simily_word, regex=True)
```





# ניתוח הנתונים – מדדי ביצוע

## Correlation Between Quality score and Service time



על פי התרשים לא נראית קורלציה בין שני משתני המטרה אך נראה פילוג לקבוצות על פי אזור שיש לבחון לעומק.



# ניתוח הנתונים – מדדי ביצוע

## בהסתכלות כללית על התפלגות הנתונים נקבעו דירוגים של: "נמוך", "בינוני", "גבוה"

### עבור איכות שירות נקבע ונצפה

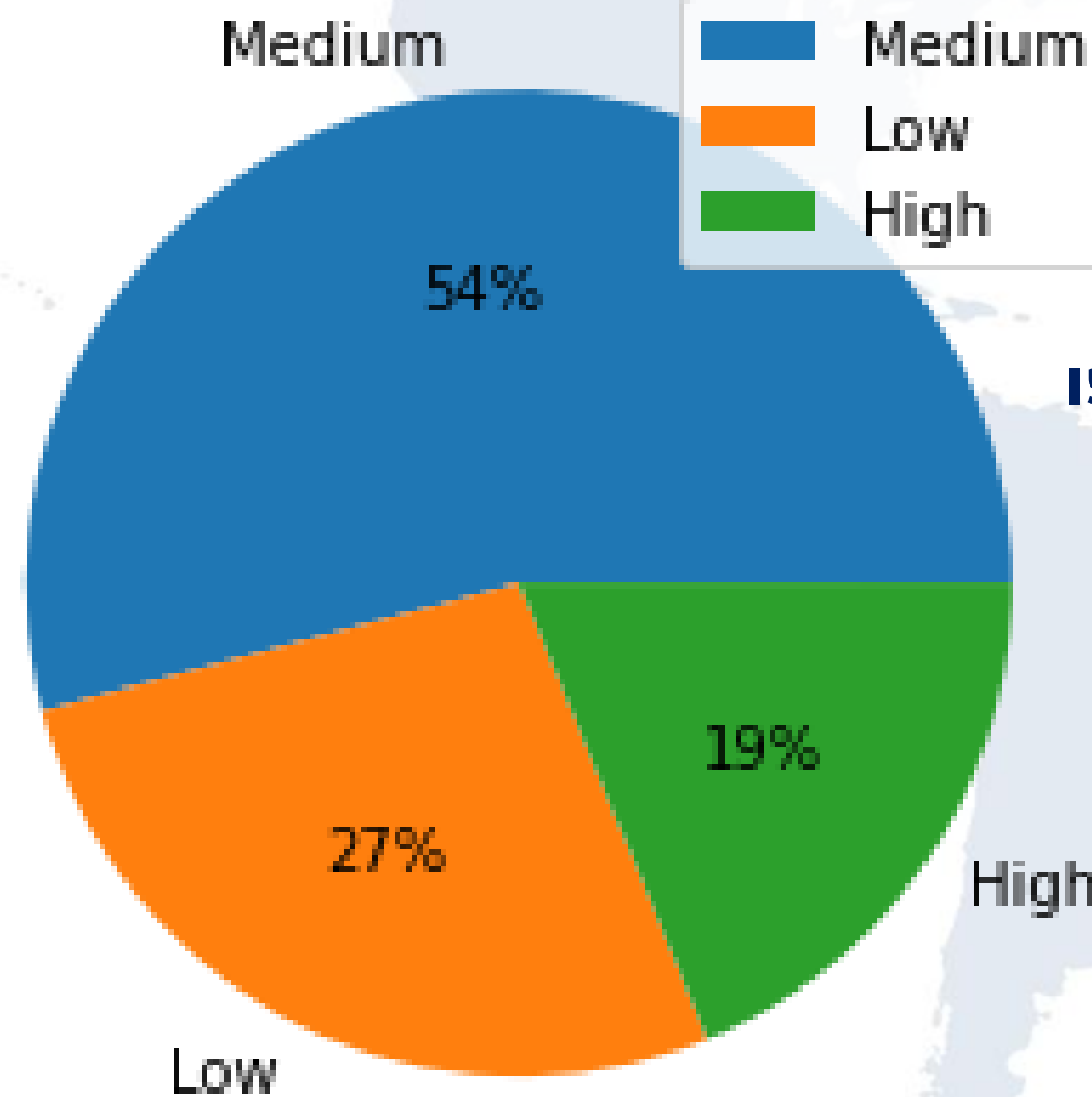
כך:

- דירוג איכות נמוך 0-6 נצפו 27% מהאירועים בחברה.

- דירוג איכות בינוני 6-8 נצפו 54% מהאירועים בחברה.

- דירוג איכות גבוה של +8 נצפו 19% מהאירועים בחברה.

ST\_Ranks

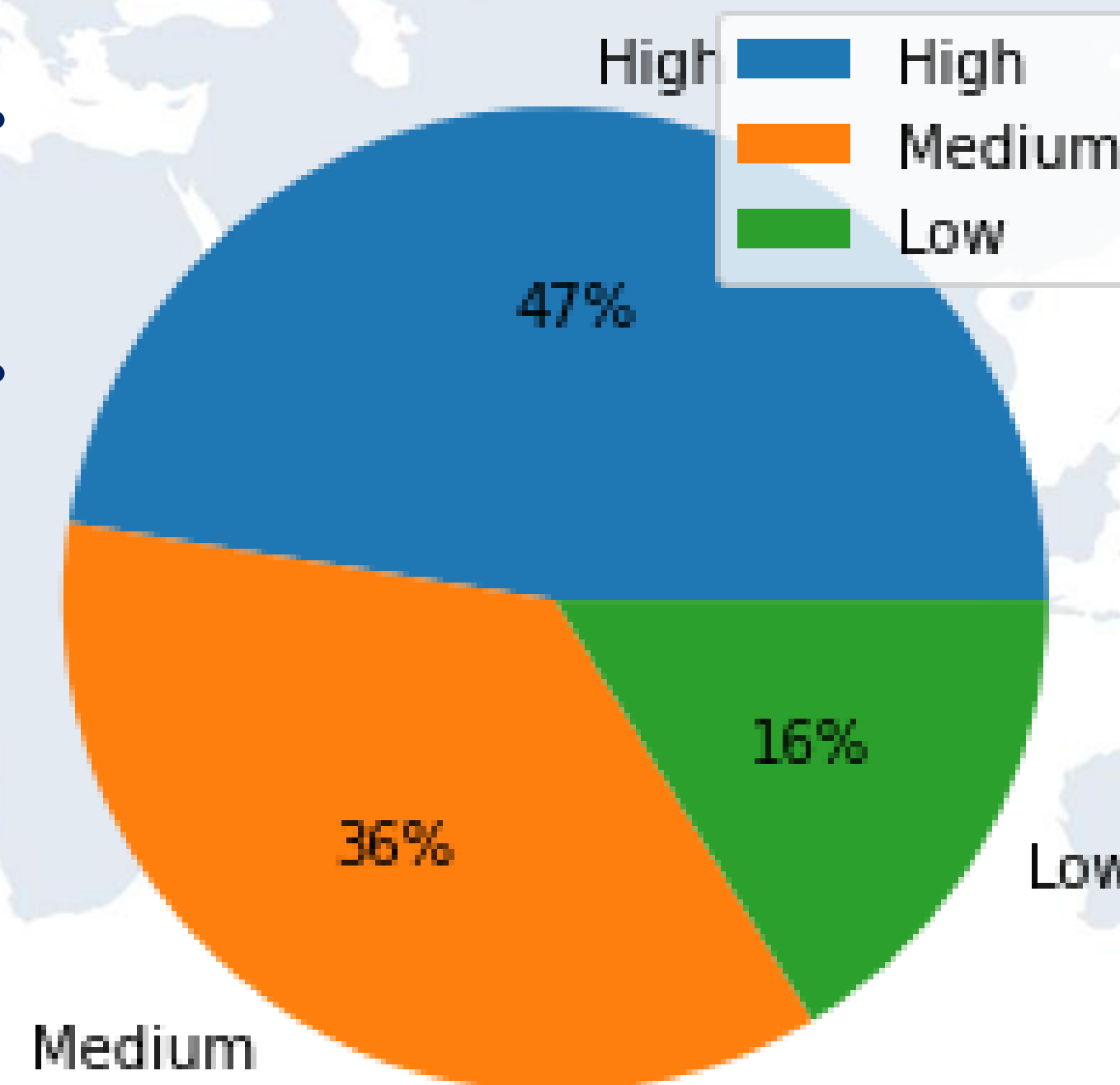


Q5\_Ranks

- זמן נמוך 0-420 זמן שירות נצפה 16% מהאירועים בחברה.

- זמן בינוני 420-470 זמן שירות נצפו 36% מהאירועים בחברה.

- זמן גבוה +470 זמן שירות נצפו 47% מהמקרים, כמעט חצי!.



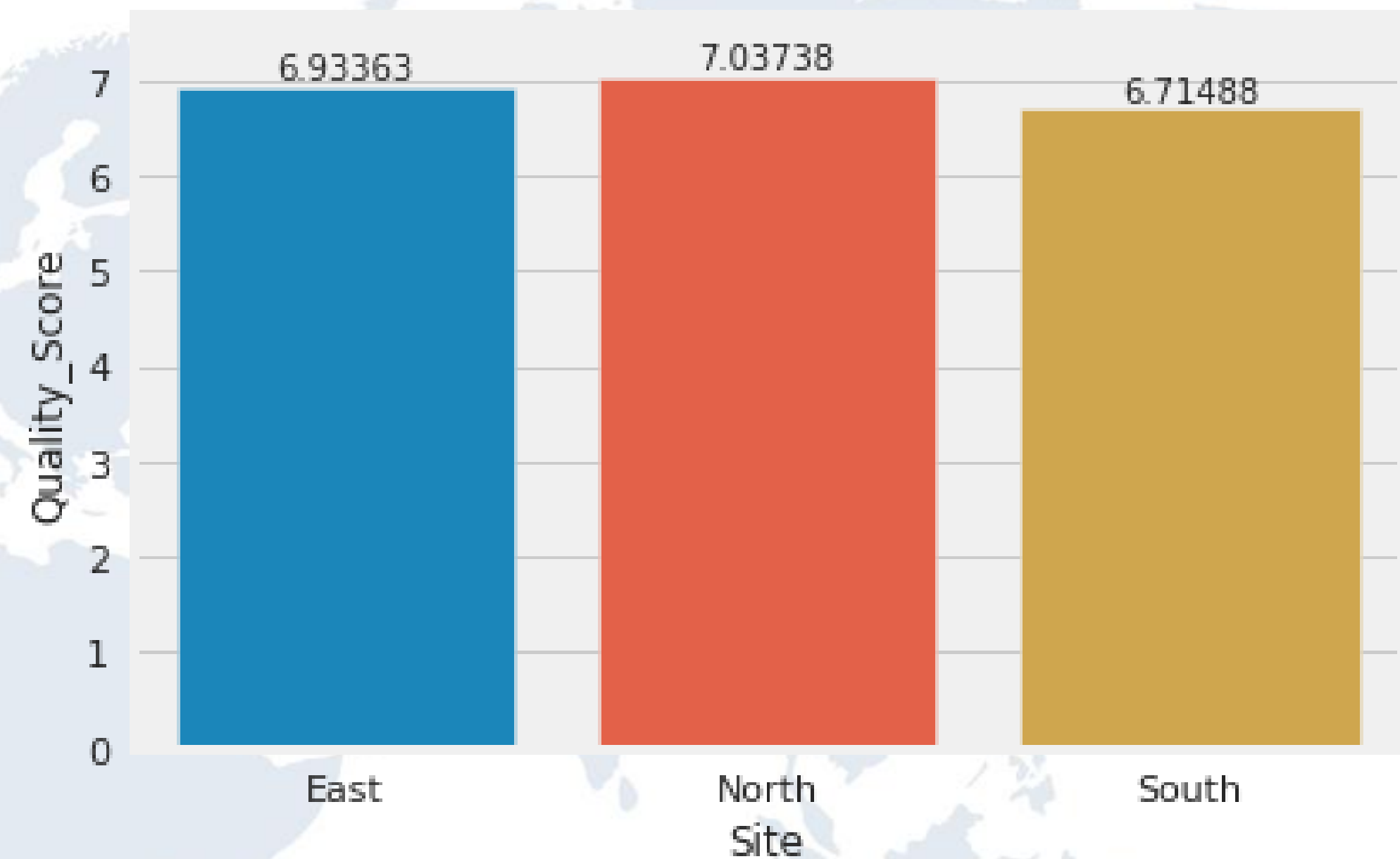
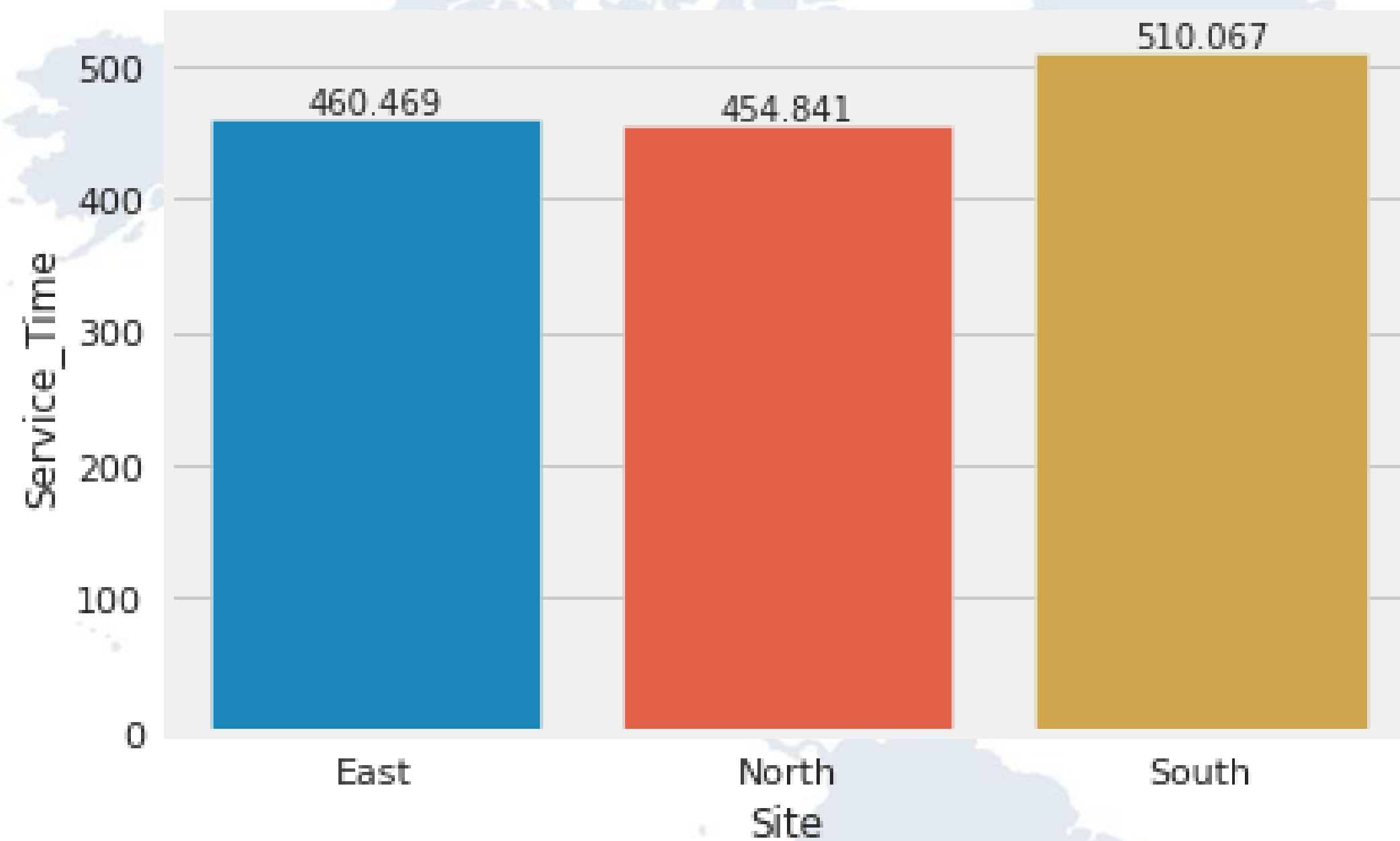
Medium

Low





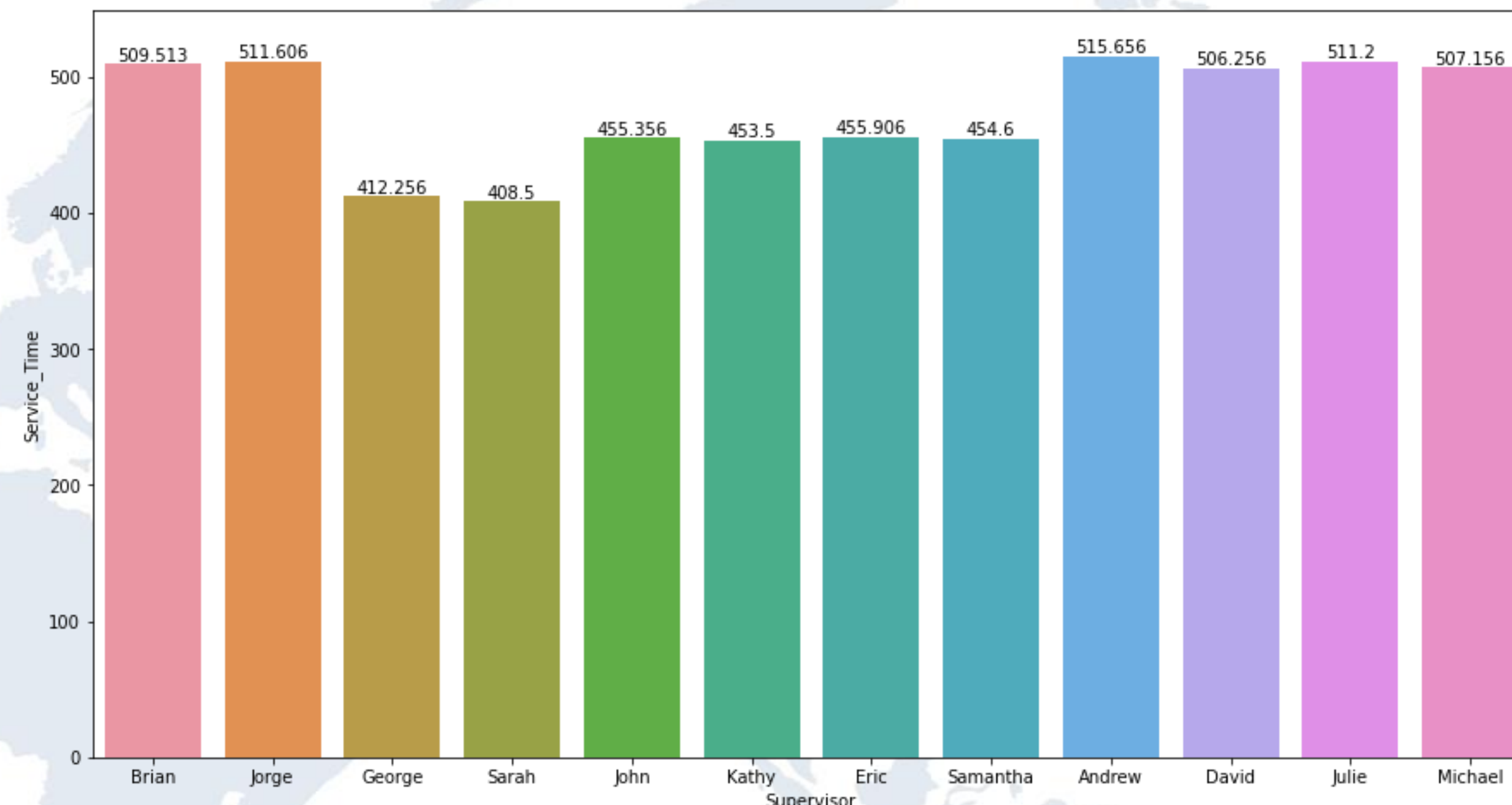
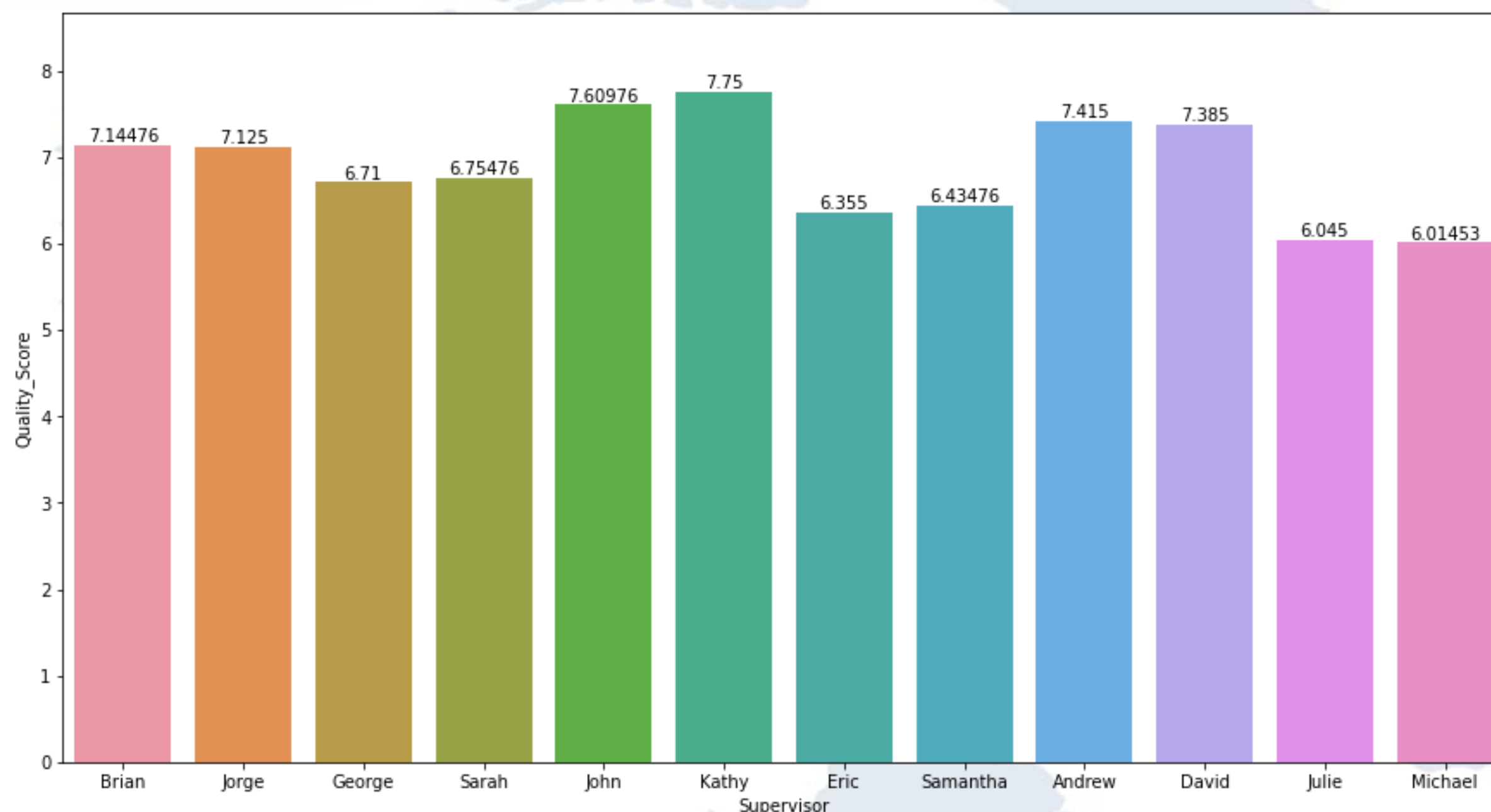
## ניתוח הנתונים - אזורים



- על פי התרשימים הבאים ניתן לראות ש:  
באזור הצפון זמן השירות הממוצע הינו הנמוך ביותר וציון האיכות הינו הגבוה ביותר.
- לעומת זאת באזור הדרום זמן השירות הממוצע הינו הגבוה ביותר וציון האיכות הממוצע הינו הנמוך ביותר.



# ניתוח הנתונים - מפקחים



**2 תרשימים אלו משווים באמצעות מדדי הביצוע "זמן שירות" ו"דירוג איכות" בין המפקחים השונים.**

**מכיוון שאנו רוצים דירוג איכות גדול ביותר אך עם זאת זמן שירות נמוך ביותר קשה להבין מן ה-2 תרשימים בנפרד האם מפקח נחשב טוב/רע ביחס לאחרים.**

**כדי על ידי מציאת מידע מועיל יותר, נשתמש בשיטות למציאת קשרים בין המדדים.**



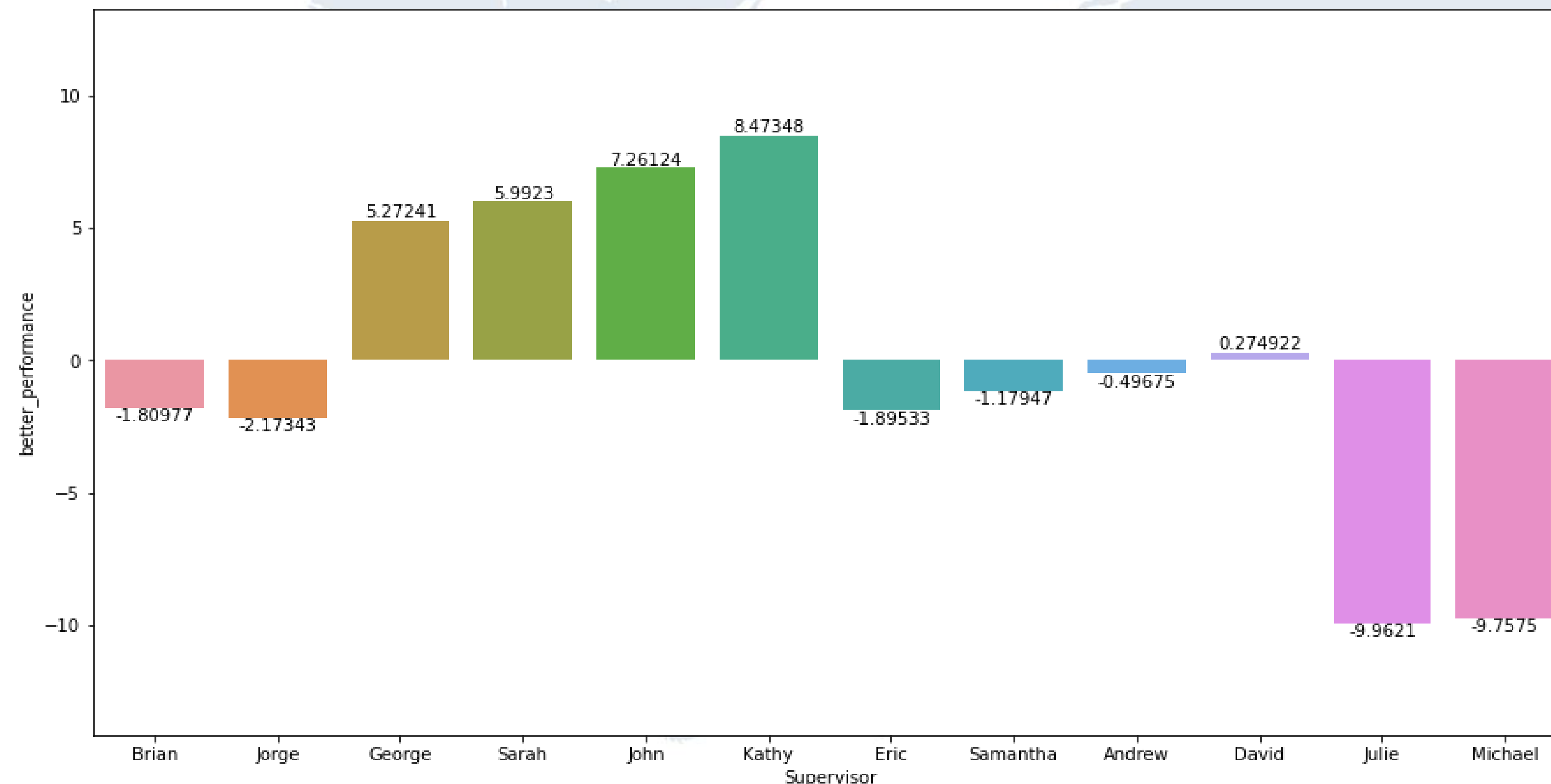


# ניתוח הנתונים – מפקחים הרחבה

הוכנסו לנתונים 3 עמודות חדשות:

1. אחוז זמן שירות שלילי – באחוזים בכמה זמן השירות שלי קטן מן הממוצע.
2. אחוז דירוג איכות חיובי – באחוזים בכמה דירוג האיכות שלי גדול מן הממוצע.
3. שירות טוב יותר - שקלול 50%-50% בין מדדי הביצוע (ניתן לשקלל שונה

בהתאם לדרישות המנהלים)



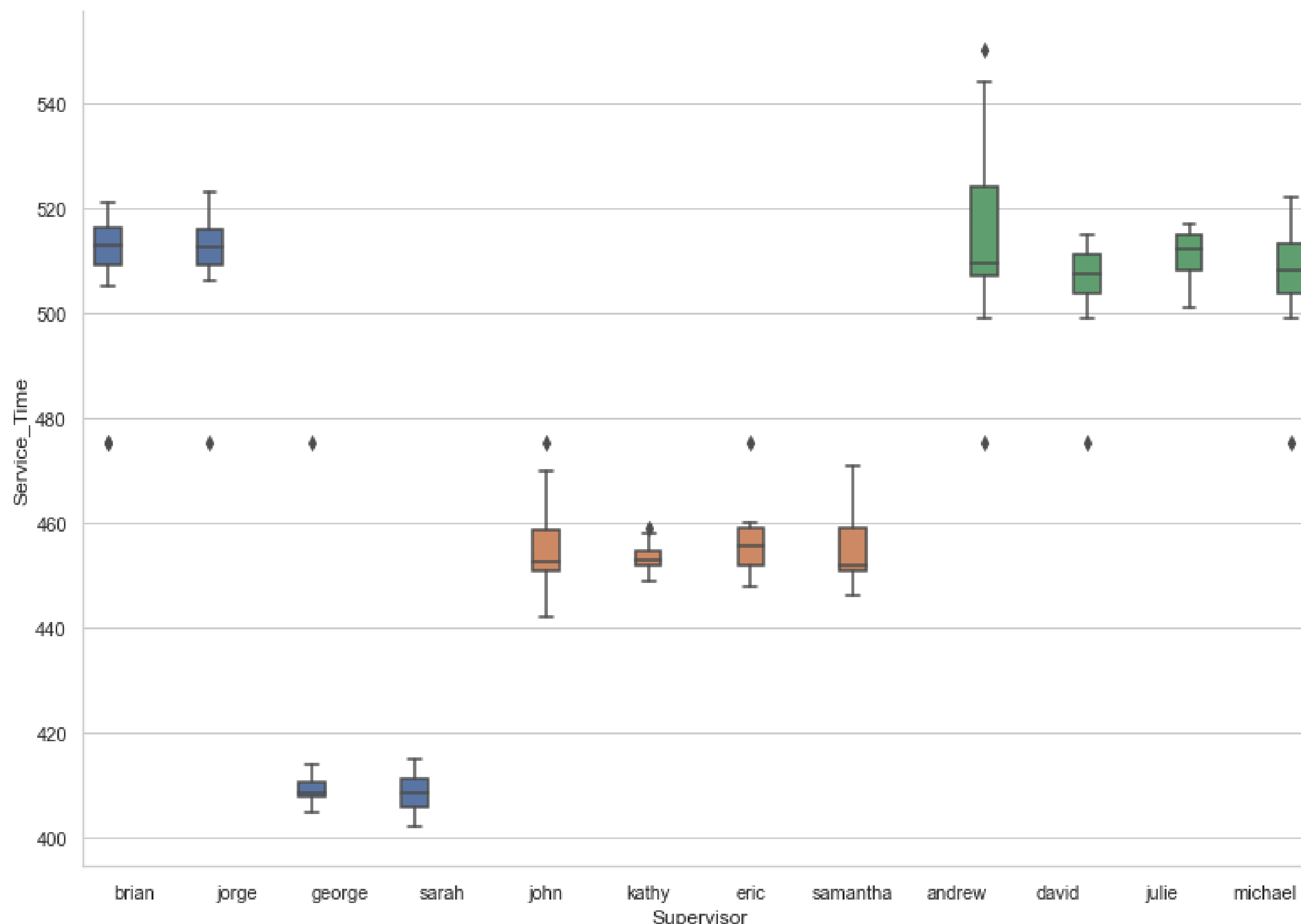
ניתן לראות את הדברים הבאים:  
מפקחת מצטיינת – Kathy עם  
אחוז ביצוע של 8.47% טוב יותר  
מהממוצע.

מפקחים שחייבים שיפור –  
Michael, Julie עם ביצועים של  
כ-9.7% נמוכים יותר מביצועי  
המפקחים בממוצע.



# ניתוח הנתונים – מפקחים הרחבה

ניתן לראות שבחברה כל 4 מפקחים על אזור ושישנם הבדלים מובהקים בזמן השירות בין האזורים.



Service_Time	
Site	
East	460.468777
North	454.840639
South	510.067208

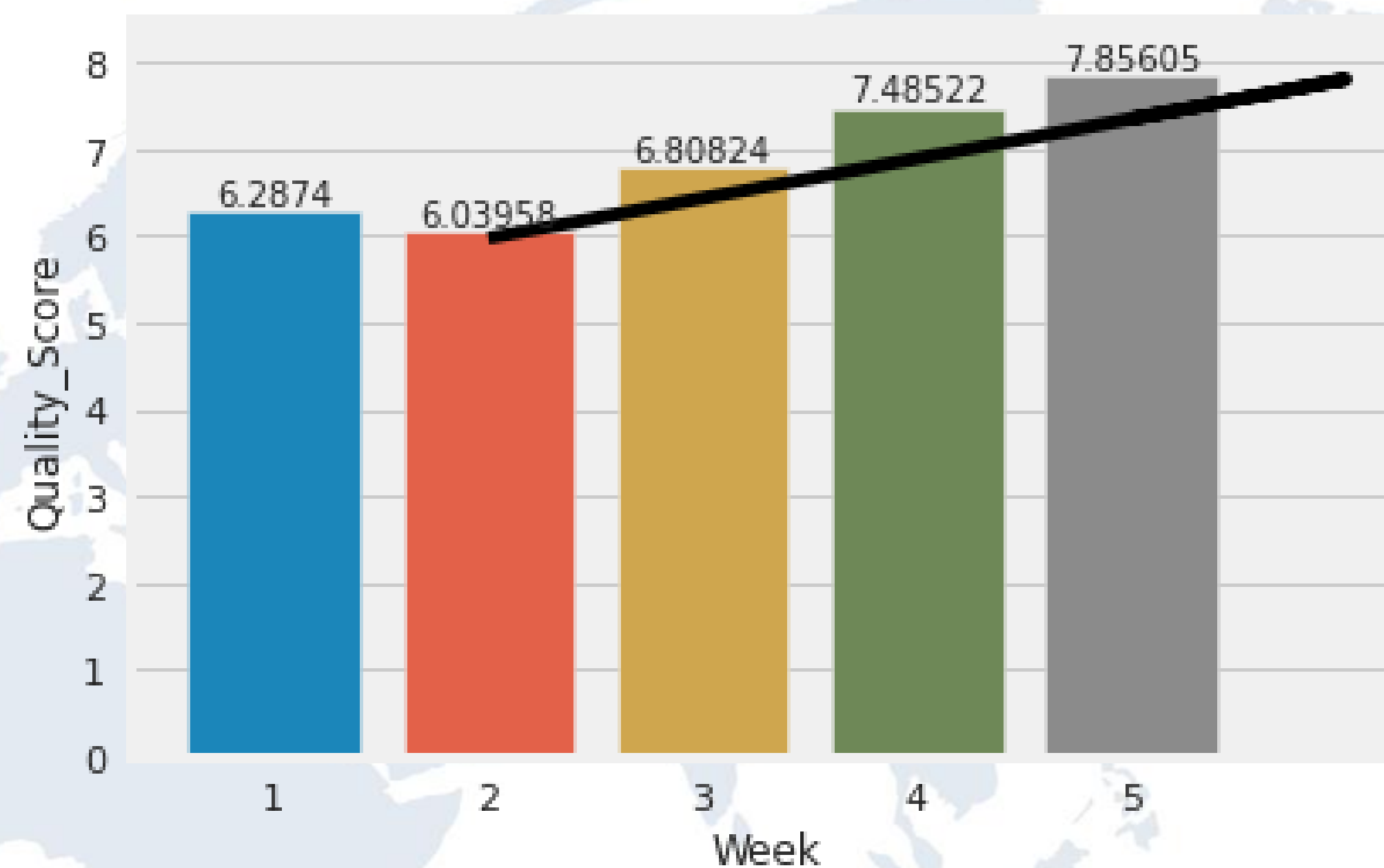
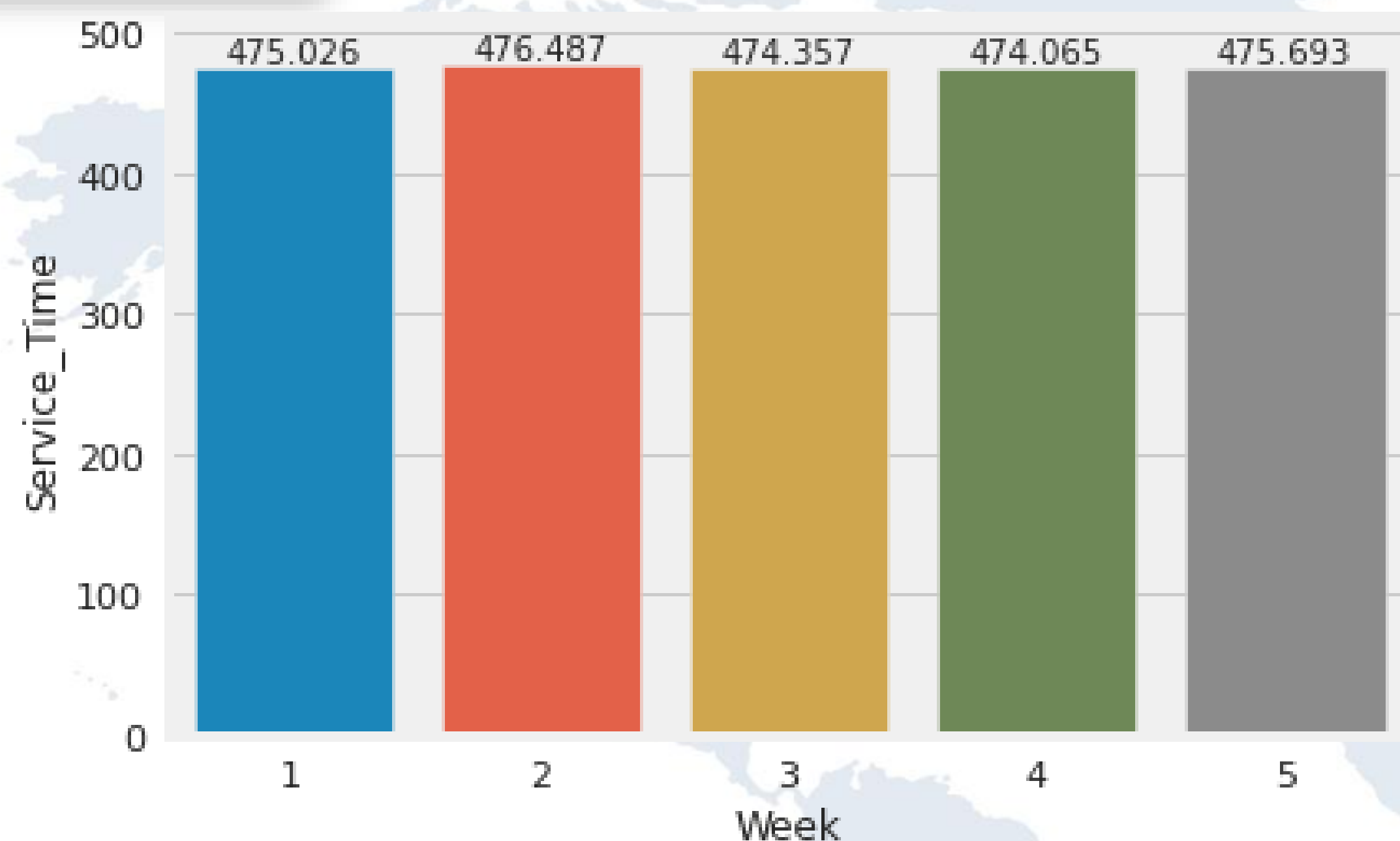
Quality_Score	
Site	
East	6.933632
North	7.037382
South	6.714882

אך עם זאת ההבדלים אינם הולכים יד ביד עם דירוג האיכות ודווקא בצפון שזמן השירות הקצר ביותר דירוג האיכות הגבוה ביותר. עלול להיות שזמן שירות ארוך מדיי מביע בעיה





## ניתוח הנתונים - שבועות



**גרפים אלו באים להציג את התפלגות מדדי הביצוע "זמן השירות" ו"דירוג האיכות" ביחס לשבועות.**

**ניתן לראות שאין הבדל מהותי בזמן השירות בין השבועות. אך עם זאת נראה שיש עלייה משמעותית עם השבועות בדירוג האיכות של השירות.**

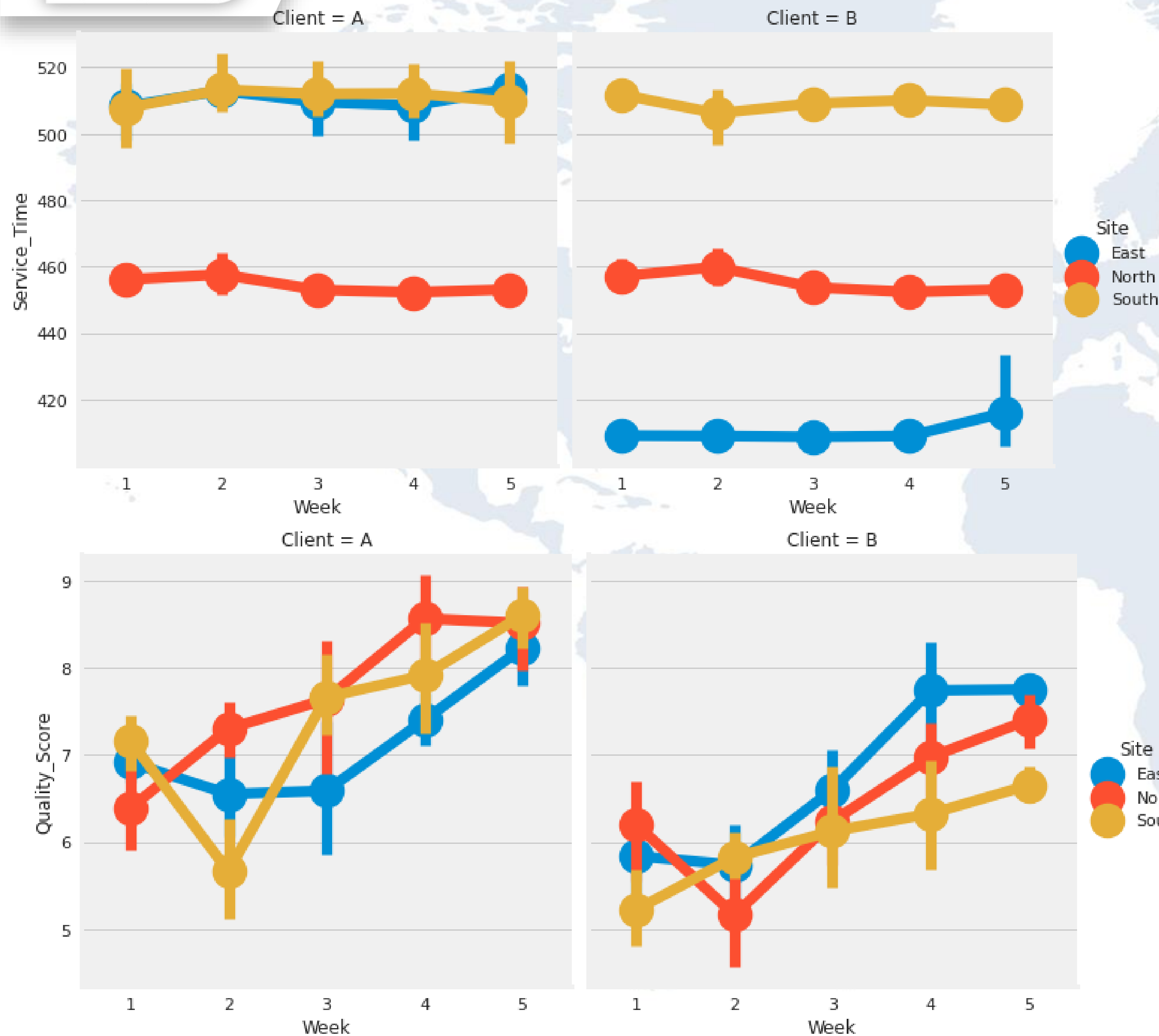


## ניתוח הנתונים – שבועות הרחבה

על מנת להבין מהיכן נובע השינוי בדירוג האיכות בין השבועות בוצעה העמקה של הניתוח יחד עם בדיקת ההבדלים בין האזורים ובין הלקוחות.

מכאן נלמד שזמן השירות במזרח ישנו הבדל גדול בין לקוח A ללקוח B.

בנוסף ניתן לראות שגם בלקוח A וגם בלקוח B עם השבועות עולה רמת איכות השירות אך עם זאת בלקוח A במזרח הוא מגיע לתוצאות נמוכות משאר האזורים ובלקוח B הוא מוביל.

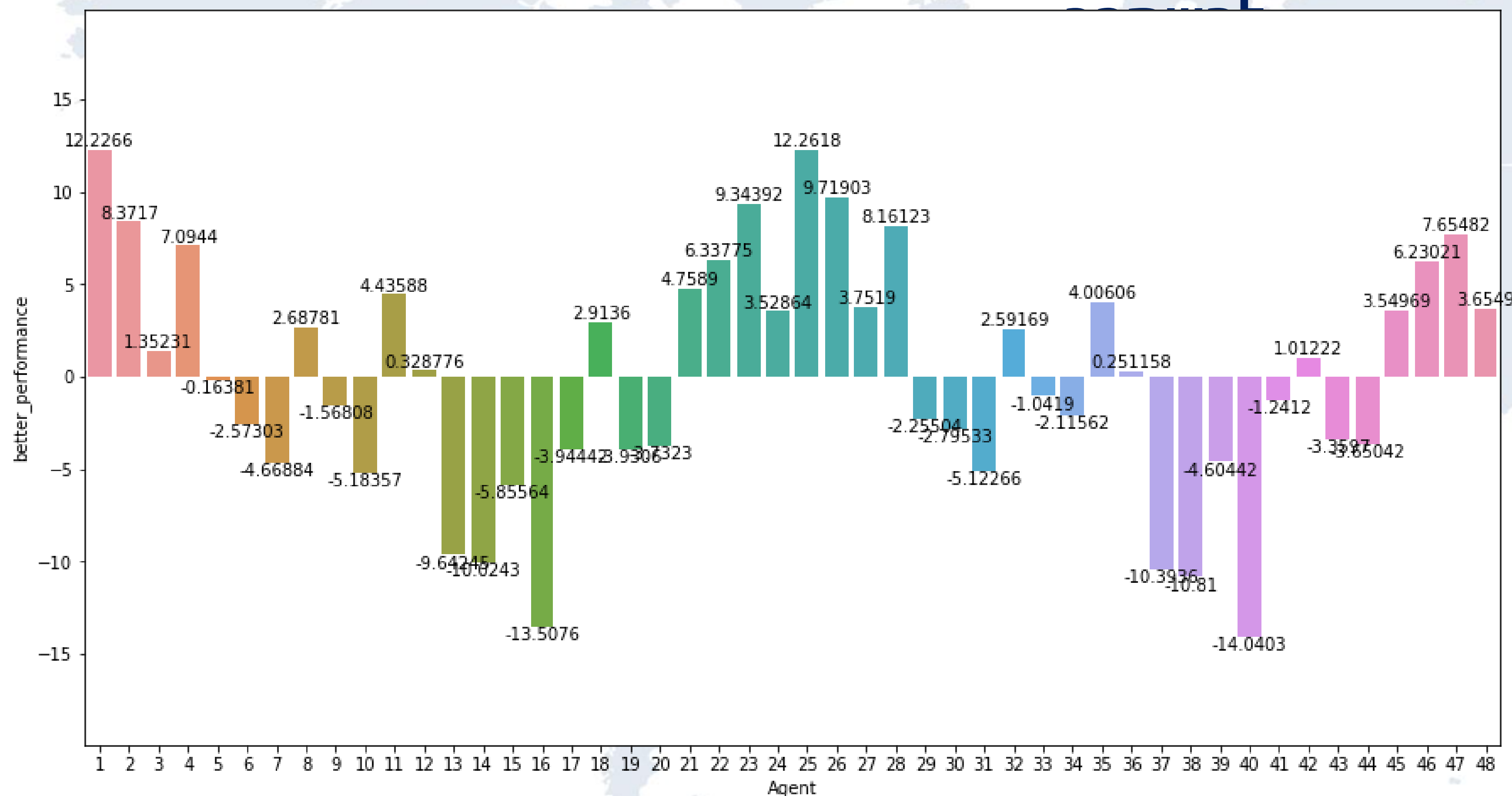






# ניתוח הנתונים – סוכנים

כמו עם המפקחים גם עם הסוכנים העובדים בחברה נרצה למדוד את הביצועים שלהם על מנת למצוא סוכנים ראויים לשבח וסוכנים שעליהם



הפעם מלכתחילה הלכנו למדד שמשקלל את 2 מדדי הביצוע 50%-50%.

ניתן לראות את הדברים הבאים:  
סוכנים מצטיינים – סוכן 1 וסוכן 25 עם ביצועים גבוהים בכ%12 מהמוצע.

סוכנים החייבים שיפור – סוכן 40 וסוכן 16 עם ביצועים נמוכים ב%13-14 מהמוצע.

\*ההנהלה יכולה להגדיר מאחוז מסוים של ביצועים לציון לשבח ומאחוז מסוים להתראה לשיחת התערבות.

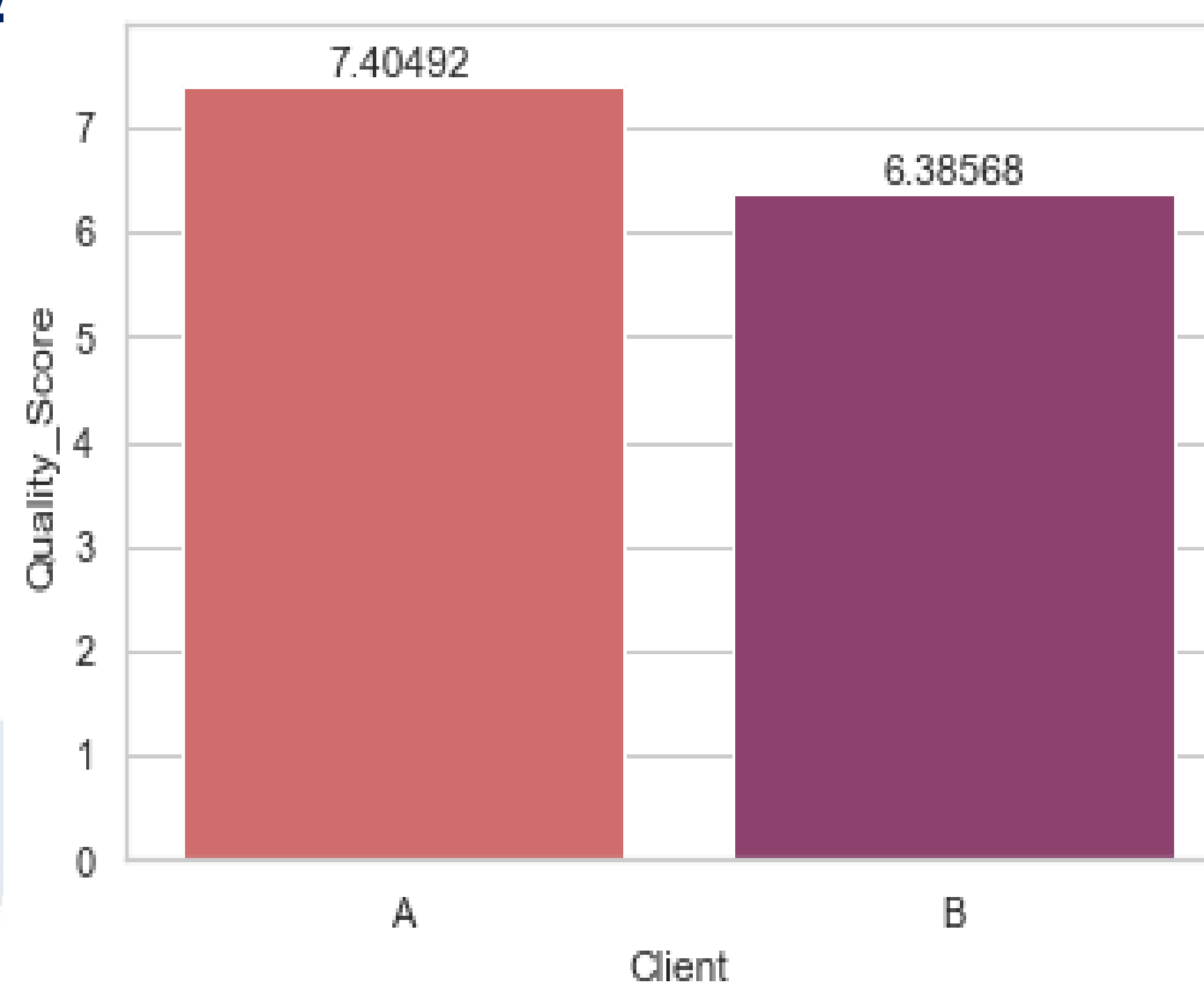
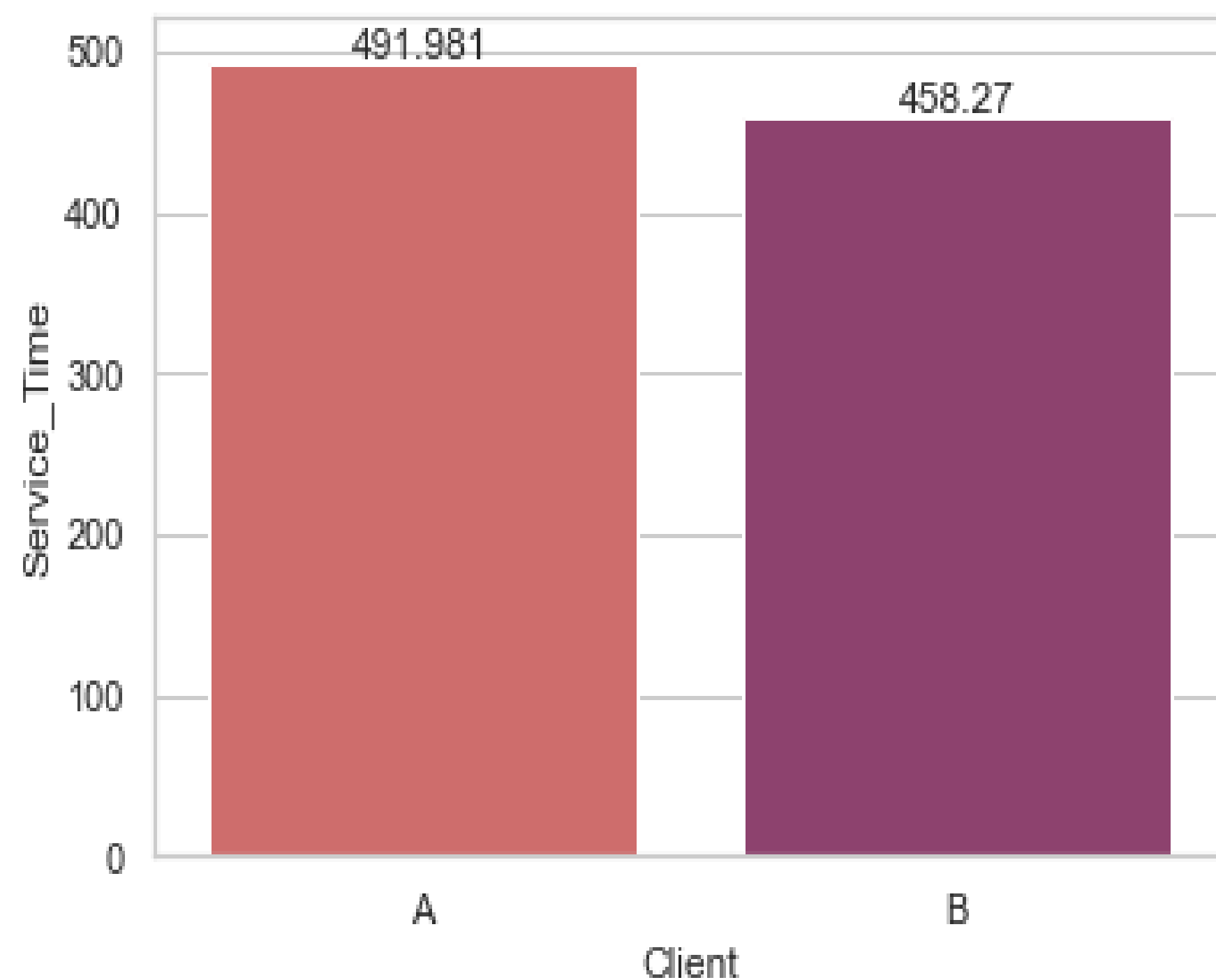


# ניתוח הנתונים – לקוחות

מתרשימים אלו ניתן לראות שלקוח A דורש יותר זמן שירות מלקוח B בכ-7%.

ניתן לראות את הפער בין לקוח A ללקוח B בדירוג האיכות וזמן השירות בין הציון

הנמוך לגבוה.



QS_Ranks	Low	Medium	High
Client			
A	18	66	36
B	47	64	9

ST_Ranks	Low	Medium	High
Client			
A	0	45	75
B	39	42	39





# סיכום ומסקנות להלן טבלה המסכמת את הממצאים הנמצאו בחקר וניתוח הנתונים:

Client	Week	Agent	Supervisor	Site	Criterion
שלקוח A דורש יותר זמן שירות מלקוח B בכ7%	שאין הבדל מהותי בזמן השירות בין השבועות.	קשה להבחין בפערים מול המפקחים ולכן נוצר מדד משוקלל המודד את היטב המפקח על פני ממוצע המפקחים בהיבט זמן השירות וציון האיכות יחדיו.	קשה להבחין בפערים מול המפקחים ולכן נוצר מדד משוקלל המודד את היטב המפקח על פני ממוצע המפקחים בהיבט זמן השירות וציון האיכות יחדיו.	<ul style="list-style-type: none"><li>באזור הצפון זמן השירות הממוצע הינו הנמוך.</li><li>באזור הדרום זמן השירות הממוצע הינו הגבוה ביותר.</li></ul>	Service Time
עם זאת דירוג האיכות מלקוח A גבוה בכ16% משל לקוח B.	יש עלייה משמעותית עם השבועות בדירוג האיכות של השירות.	המלצה להנהלה – לבחון גבול עליון ותחתון לבקרת הסוכנים.		<ul style="list-style-type: none"><li>באזור הצפון ציון האיכות הממוצע הגבוהה ביותר.</li><li>באזור הדרום ציון האיכות הממוצע המוך ביותר.</li></ul>	Quality Score
ההבדל נובע בין היתר מההבדל בין המפקחים והסוכנים. ייתכן שבנוסף יש להקפיד על עלייה בהבדלים נובעים מצרכי לקוח שונים.	על פני השבועות נראה שיפור הולך ועולה בביצועים הכללים זאת בעיקר מעלייה בדירוג האיכות.	סוכנים מצטיינים – סוכן 1 וסוכן 25 עם ביצועים גבוהים בכ12% מהממוצע. סוכנים החייבים שיפור – סוכן 40 וסוכן 16 עם ביצועים נמוכים ב13%-14% מהממוצע.	מפקחת מצטיינת – Kathy עם אחוז ביצוע של 8.47% טוב יותר מהממוצע. מפקחים שחייבים שיפור – Michael, Julie עם ביצועים של כ9.7% נמוכים יותר מביצועי המפקחים בממוצע. *לשים לב להכנסה נכונה של שם המפקח במערכת.	ישנם הבדלים בזמן השירות בין האזורים ועל פי ניתוח מעמיק ייתכן והשינוי ניכר בגלל סוג הלקוח.	General

**ARIEL**  
**UNIVERSITY**

אריאל  
אוניברסיטה



**תודה רבה**

**מחויבים להצלחה שלך**

**מייל לשאלות נוספות – [liadby1105@gmail.com](mailto:liadby1105@gmail.com)**