

Challenge Instructions

Your client is a broker for orthopedic surgeries who wants to understand the performance of their surgeons. He is asking you to analyze his data and answer the following question:

Who are the most and least skilful surgeons for hip replacement operations?

Make sure to include all your work in your email when sending us your answers. You are welcome to use tools of your choice when completing this task, so long as you send all relevant documents together with your answer (for example, include the SQL queries if you are using SQL to analyze the data). Be sure to carefully study the background information.

Background Information

- The client works with data from the EQ-5D-5L questionnaire - a survey that is taken by patients both before and after their operations to measure outcome success. By comparing the scores, the improvement in how patients feel in various areas of their lives can be evaluated.
- The questionnaire consists of five sections: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.
- The success of the surgery is evaluated by comparing the quality of life pre and post-operation.
- The answers will then be converted to a health state score. Please see below for an example questionnaire and an example of conversion.
- Based on the example questionnaire (on the next page), Patient A has reported their symptoms to be the following:
 - Mobility: Slight problems (“Slight”)
 - Self-care: Moderate problems (“Moderate”)
 - Usual activities: Slight problems (“Slight”)
 - Pain/discomfort: Severe problems (“Severe”)
 - Anxiety/depression: Extreme problems (“Unable”)
- The value set you will need for the conversion from individual answers to a health state score is part of the data set, in the table “answer_options”. Each answer has a corresponding “central estimate”. The “None” category is always 0, therefore it was omitted from the table.
- In order to receive a score for the patient’s health state, it is necessary to deduct the sum of the individual scores from 1, as can be seen in the table above. Therefore, a higher score means better health. In our example, the overall health score would be 0.247, calculated as follows:

1 - (0.058 + 0.080 + 0.050 + 0.276 + 0.289) = 0.247

Data Understanding

| | question_id int64 1 - 5 | answer object I have no pro... 4% I have slight ... 4% 23 others 92% | severity_code int64 1 - 5 | central_estim 0.0 - 0.335 |
|---|----------------------------|---|------------------------------|------------------------------|
| 0 | 1 | I have no problem... | 1 | |
| 1 | 1 | I have slight probl... | 2 | |
| 2 | 1 | I have moderate ... | 3 | |
| 3 | 1 | I have severe pro... | 4 | |
| 4 | 1 | I am unable to wa... | 5 | |

| | id int64 183646 - 183655 | question_id int64 1 - 5 | patient_id int64 20959 - 20959 | questionnaire |
|--------|-----------------------------|----------------------------|-----------------------------------|---------------|
| 183... | 183646 | 1 | 20959 | |
| 183... | 183647 | 2 | 20959 | |
| 183... | 183648 | 3 | 20959 | |
| 183... | 183649 | 4 | 20959 | |
| 183... | 183650 | 5 | 20959 | |
| 183... | 183651 | 1 | 20959 | |
| 183... | 183652 | 2 | 20959 | |
| 183... | 183653 | 3 | 20959 | |
| 183... | 183654 | 4 | 20959 | |
| 183... | 183655 | 5 | 20959 | |

| | id int64 1 - 10 | name object Princess Leia 10% Boba Fett 10% 8 others 80% | |
|---|--------------------|---|--|
| 0 | 1 | Princess Leia | |
| 1 | 2 | Boba Fett | |
| 2 | 3 | Darth Vader | |
| 3 | 4 | Han Solo | |
| 4 | 5 | Obi-Wan Kenobi | |

| | id int64 | title object | desc |
|---|----------|----------------------|------|
| 0 | 1 | Mobility | Plea |
| 1 | 2 | Self-care | Plea |
| 2 | 3 | Usual activities | Plea |
| 3 | 4 | Pain / discomfort | Plea |
| 4 | 5 | Anxiety / depress... | Plea |

| | id int64 1 - 25000 | gender object Female 50.3% Male 49.7% | surg 1 - 10 |
|---|-----------------------|---|----------------|
| 0 | 1 | Male | |
| 1 | 2 | Male | |
| 2 | 3 | Female | |
| 3 | 4 | Male | |
| 4 | 5 | Female | |

Data preparation

- Data cleaning: Is the process of fixing or removing incorrect, corrupted, duplicate, redundante or incomplete data within a dataset.
- Metrics calculation: To keep the final process simple and facilitate the analysis, the metrics are already calculate excluding also the fields that become redundante

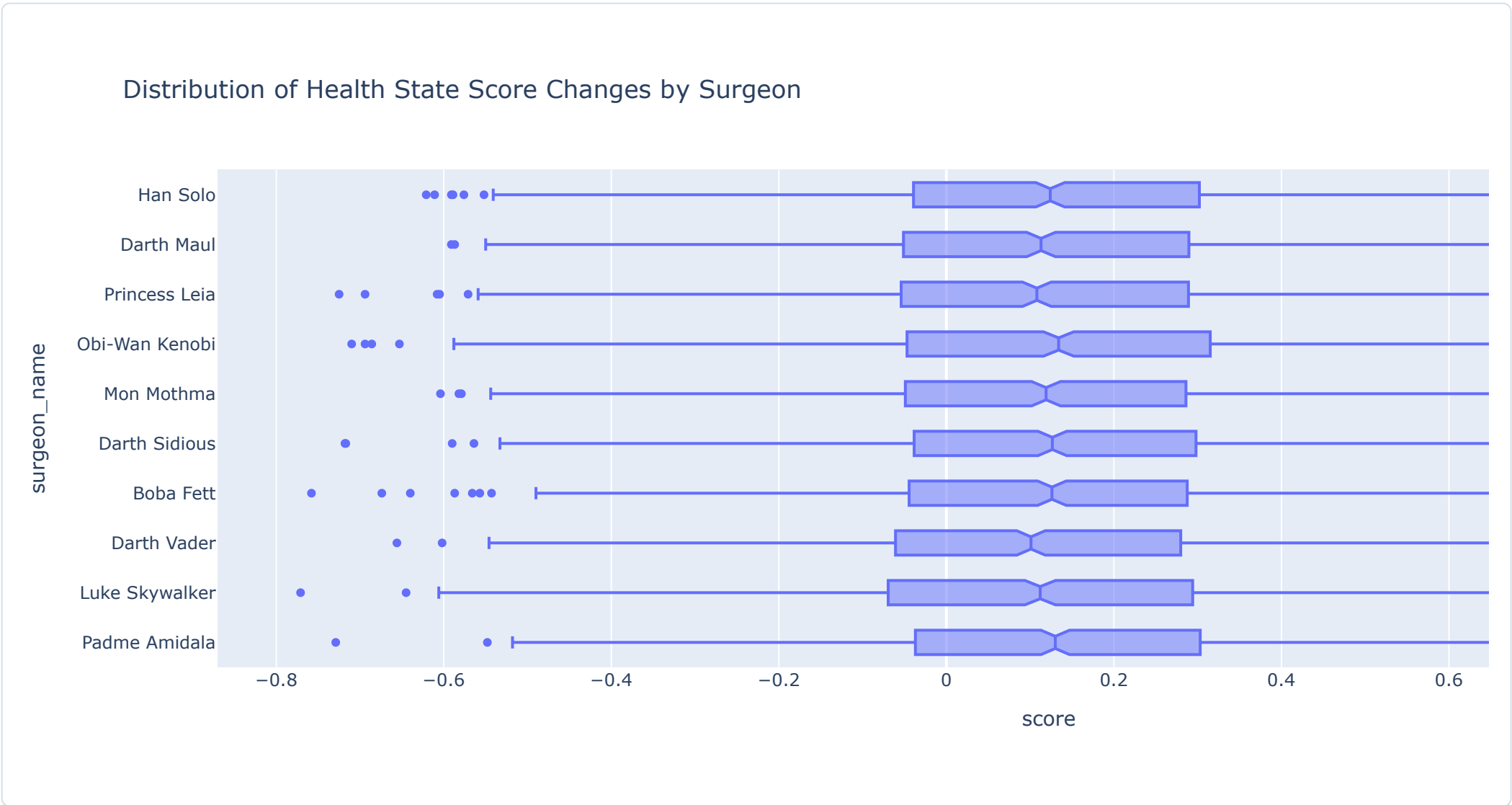
```
WITH base AS (  
  SELECT *  
    , CASE WHEN questionnaire_id = 1 THEN 'pre' ELSE 'post'  
  -- Check if there is data for pre and post  
    , MAX(questionnaire_id) OVER(PARTITION BY patient_id) AS max_id  
    , MIN(questionnaire_id) OVER(PARTITION BY patient_id) AS min_id  
  FROM answers  
  WHERE questionnaire_id in (1,2) -- Remove the Knee treatment  
)  
 , pre_post AS (  
  SELECT id, question_id, patient_id, answer, type  
  FROM base  
  WHERE max_id = 2 AND min_id = 1 -- Select only patients that have both pre and post  
)  
  
SELECT  
  pre_post.*, code.severity_code, code.central_estimate  
FROM pre_post  
LEFT JOIN answer_options AS code USING(question_id, answer)
```

| | id int64 1 - 219080 | question_id int64 | patient_id int64 | answer object |
|---|------------------------|-------------------|------------------|---------------|
| 0 | 222 | 2 | 26 | I have no pro |
| 1 | 223 | 3 | 26 | I have severe |
| 2 | 224 | 4 | 26 | I have moder |
| 3 | 225 | 5 | 26 | I am severely |
| 4 | 226 | 1 | 26 | I have no pro |

```
-- Scores calculation  
  
SELECT  
  name AS surgeon_name  
  , patient_id  
  , 1 - SUM(CASE WHEN type = 'pre' THEN central_estimate END)  
  , 1 - SUM(CASE WHEN type = 'post' THEN central_estimate END)  
  , post - pre as score  
FROM answer_reduc AS h  
LEFT JOIN patients AS p ON h.patient_id = p.id  
LEFT JOIN surgeons AS s ON s.id = p.surgeon_id  
  
GROUP BY 1,2
```

| | surgeon_name o... | patient_id int64 | pre float64 | post float64 |
|---|--|------------------|---------------------|---------------|
| | Darth Vader 10.3% Obi-Wan Ke... 10.2% 8 others 79.5% | 1 - 25000 | -0.2809999999999... | -0.1999999999 |
| 0 | Padme Amidala | 67 | 0.382 | |
| 1 | Luke Skywalker | 118 | 0.391 | |
| 2 | Darth Vader | 134 | 0.388 | |
| 3 | Boba Fett | 280 | 0.449 | |
| 4 | Padme Amidala | 301 | 0.51 | |

```
# Group the data by surgeon and calculate the average score difference and standard deviation  
  
surgeon_performance = health_state.groupby('surgeon_name').score.agg(  
    avg_score='mean',  
    std_score='std',  
    min_score='min',  
    max_score='max'  
) .reset_index()  
  
surgeon_performance['negative_improvement_rate'] = health_state.groupby('surgeon_name').apply(  
    lambda x: (x['score'] < 0).mean()  
) .values  
  
surgeon_performance = surgeon_performance.sort_values(  
    by=['avg_score', 'std_score', 'negative_improvement_rate'],  
    ascending=[False, True, True]  
) .reset_index(drop=True)
```



Who are the most and least skillful surgeons for hip replacement operations?

In a glance: The global most skillful surgeons for hip replacement are Padme Amidala, Obi-Wan Kenobi, and Darth Sidious. Conversely, the least skillful surgeons are Luke Skywalker, Darth Vader, and Mon Mothma.

To refine the evaluation, some criteria are considered to reflect not only improving outcomes on average but also delivering consistent and safer results.

Average Score:

- Top Surgeons:
- 1- Padme Amidala with an average score of 0.132
 - 2- Obi-Wan Kenobi with an average score of 0.131
 - 3- Darth Sidious with an average score of 0.124
- Bottom Surgeons:
- 1- Luke Skywalker with an average score of 0.111
 - 2- Darth Vader with an average score of 0.112
 - 3- Mon Mothma with an average score of 0.113

Notably, Padme Amidala exhibits not only a high average score but also a low negative improvement rate and the highest maximum score, signifying consistent excellence. Conversely, Luke Skywalker's low average score, coupled with the highest negative improvement rate and lowest maximum score, suggests consistent underperformance.

Standard Deviation of Scores (std_score):

- Lower standard deviations indicate greater consistency in performance. Surgeons like Mon Mothma and Darth Vader exhibit the least variability, indicating a more predictable performance pattern. Conversely, higher-ranked surgeons such as Padme Amidala and Obi-Wan Kenobi display higher deviations, implying a broader spectrum of outcomes ranging from peaks to troughs.
- Top Surgeons (Lowest rate):
- 1- Mon Mothma with a standard deviation of 0.251
 - 2- Darth Vader with a standard deviation of 0.252
 - 3- Princess Leia with a standard deviation of 0.252
- Bottom Surgeons (Highest rate):
- 1- Luke Skywalker with a standard deviation of 0.264
 - 2- Padme Amidala with a standard deviation of 0.261
 - 3- Obi-Wan Kenobi with a standard deviation of 0.260

Negative Improvement Rate:

- Measures the frequency of performance declines. Obi-Wan Kenobi has the lowest rate, suggesting fewer instances where performance worsened. The same three surgeons that are in the top three of the average score are also the ones with the lowest negative improvement rates.
- Top Surgeons (Lowest rate):
- 1- Obi-Wan Kenobi with a negative improvement rate of 0.297
 - 2- Padme Amidala with a negative improvement rate of 0.301
 - 3- Darth Sidious with a negative improvement rate of 0.305
- Bottom Surgeons (Highest rate):
- 1- Luke Skywalker with a negative improvement rate of 0.342
 - 2- Princess Leia with a negative improvement rate of 0.334
 - 3- Darth Maul with a negative improvement rate of 0.325

Conclusion:

The analysis underscores the importance of considering both average outcomes and consistency when evaluating surgeons' performance. While high average scores indicate proficiency, factors such as standard deviation and negative improvement rates provide valuable insights into reliability and safety. Surgeons like Padme Amidala and Obi-Wan Kenobi emerge as exemplary performers, showcasing not only superior skill but also consistent and reliable outcomes. Conversely, surgeons like Luke Skywalker exhibit consistent underperformance, highlighting the significance of addressing variability and enhancing skill proficiency for improved surgical outcomes.

For a clearer visual representation, refer to the chart below. Here, the highest points correspond to higher average scores, while points horizontally closer from the origin indicate a lower rate of performance declines. Ideally, the top-performing surgeons should cluster on the top, near the y-axis for optimal positioning.

