


## Healthcare Project - Profitability of Various Procedures

We calculate the total revenue generated by each procedure and assess its profitability by comparing the revenue to associated costs. This helps us identify the most financially efficient procedures, prioritize high-margin services, and pinpoint areas that may require cost optimization or reevaluation.

```
# Modules
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt


# Load data
path = "C:/Users/rvrei/Documents/Healthcare_df.csv"
healthcare_df = pd.read_csv(path)
healthcare_df.head(3)
```



	claim_id	patient_id	procedure_id	claim_date	claim_amount	claim_status	insurance_provider	procedure_type	pr
0	CLM0001	PAT0001	41	2024-03-29	1997.79	approved	Blue Shield	CT Scan	
1	CLM0001	PAT0001	41	2024-03-29	1997.79	approved	Blue Shield	CT Scan	
2	CLM0016	PAT0016	41	2023-09-16	1080.34	approved	Aetna	MRI	

3 rows × 21 columns

```
# Calculate profitability for each procedure: revenue - cost
healthcare_df['profit'] = healthcare_df['revenue'] - healthcare_df['procedure_cost']
healthcare_df.head(3)
```



	claim_id	patient_id	procedure_id	claim_date	claim_amount	claim_status	insurance_provider	procedure_type	pr
--	----------	------------	--------------	------------	--------------	--------------	--------------------	----------------	----

0	CLM0001	PAT0001	41	2024-03-29	1997.79	approved	Blue Shield	CT Scan
1	CLM0001	PAT0001	41	2024-03-29	1997.79	approved	Blue Shield	CT Scan
2	CLM0016	PAT0016	41	2023-09-16	1080.34	approved	Aetna	MRI

3 rows × 22 columns

```
# # Calculate total Revenue, total Cost, and total Profit
print(f"Total Revenue: ${round(healthcare_df['revenue'].sum(),2):,}")
print(f"Total Cost: ${round(healthcare_df['procedure_cost'].sum(),2):,}")
print(f"Total Profit: ${round(healthcare_df['profit'].sum(), 2):,}")
```

```
Total Revenue: $660,531.79
Total Cost: $387,519.51
Total Profit: $273,012.28
```

```
# Group by procedure_name and calculate total profit
profitability_df = healthcare_df.groupby('procedure_type').agg(total_revenue=('revenue','sum'),\
    total_cost=('procedure_cost','sum'),total_profit=('profit','sum')).reset_index()
profitability_df = profitability_df.sort_values(by='total_profit', ascending=False)
```

```
profitability_df['total revenue ($)'] = profitability_df['total_revenue'].apply(lambda x: f"${x:,.2f}")
profitability_df['total cost ($)'] = profitability_df['total_cost'].apply(lambda x: f"${x:,.2f}")
profitability_df['total profit ($)'] = profitability_df['total_profit'].apply(lambda x: f"${x:,.2f}")
profitability_df.drop(['total_revenue','total_cost','total_profit'], axis=1)
```

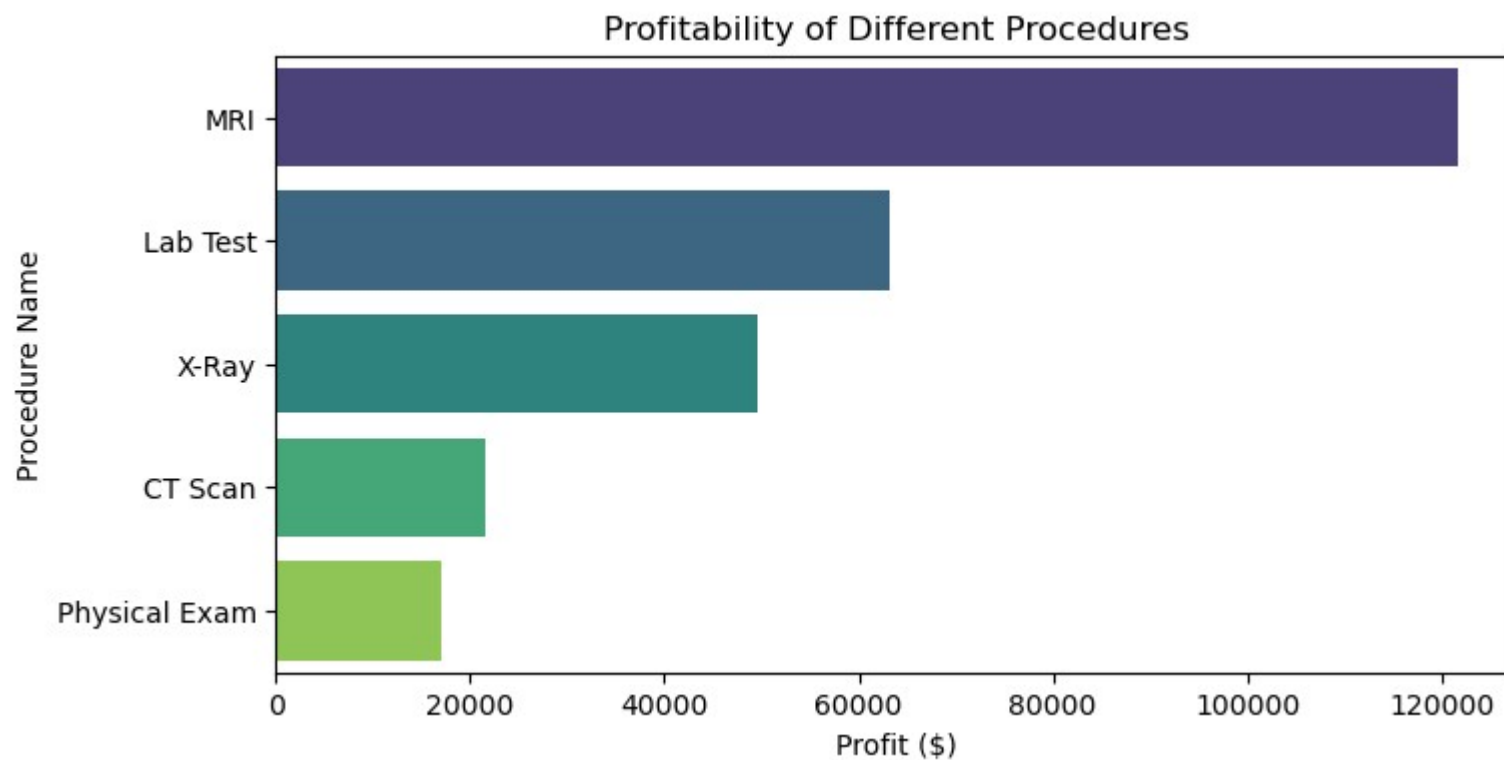
```


```

	procedure_type	total revenue (\$)	total cost (\$)	total profit (\$)
2	MRI	\$272,910.28	\$151,208.13	\$121,702.15
1	Lab Test	\$146,980.83	\$83,787.66	\$63,193.17
4	X-Ray	\$111,395.57	\$61,889.94	\$49,505.63
0	CT Scan	\$55,699.21	\$34,156.26	\$21,542.95

3	Physical Exam	\$73,545.90	\$56,477.52	\$17,068.38
---	---------------	-------------	-------------	-------------

```
# Visualization: Plot Profitability by Procedure
plt.figure(figsize=(8,4))
sns.barplot(x='total_profit', y='procedure_type', data=profitability_df, palette='viridis')
plt.title('Profitability of Different Procedures')
plt.xlabel('Profit ($)')
plt.ylabel('Procedure Name')
plt.show()
```



This breakdown offers insight into the profitability of each procedure, with MRI being the most profitable and the Physical Exam the least. This information is valuable for assessing which services contribute most to the organization's overall profitability.

