

Healthcare Analysis

1. Project Overview

This project analyzes hospital patient flow using Healthcare analytics patient flow data with 9,216+ patient encounter records from Sept 2023–Dec 2024. The goal is to uncover into actionable insights of demographics (age, gender, race), arrival time patterns, department segments, admission decisions, wait times, and satisfaction scores. To guide business decisions.

2. Summary

-Rows: 9,216

-Columns: 11

-Features:

-Patient demographics age, gender, race

-Admission details Patient Admission Date/Time, Patient, Department Referral, Patient Admission Flag ,Patient Wait time, Patient Satisfaction Score.

- Missing data: 5401 values in Department Referral column, 6699 values in Patient Satisfaction Score

3. Exploratory Data Analysis using Python

❖ DATA loading: Imported dataset using pandas

	Patient Id	Patient Admission Date	Patient Admission Time	Merged	Patient Gender	Patient Age	Patient Race	Department Referral	Patient Admission Flag	Patient Satisfaction Score	Patient Waittime
0	780-96-6113	9/9/2024	9:25:00 AM	W. Breede	Female	63	African American	NaN	Not Admission	5.0	32
1	714-35-6722	9/9/2024	4:42:00 PM	Y. Baldetti	Male	31	Asian	Orthopedics	Not Admission	NaN	22
2	571-85-3714	9/9/2024	12:14:00 AM	M. Semerad	Male	75	White	General Practice	Not Admission	NaN	16
3	404-43-9499	9/9/2024	8:33:00 PM	K. Blaydes	Male	79	African American	General Practice	Admission	NaN	38
4	552-51-5855	9/9/2024	7:25:00 PM	F. Dickerson	Female	24	African American	NaN	Admission	NaN	36
...
9211	353-25-3957	1/1/2024	11:01:00 PM	S. Levermore	Male	19	White	Cardiology	Not Admission	NaN	34
9212	608-47-0337	1/1/2024	5:25:00 PM	R. Boyle	Female	22	White	Orthopedics	Admission	NaN	45
9213	716-90-5923	1/1/2024	11:42:00 AM	V. Seardon	Female	8	African American	NaN	Admission	NaN	21
9214	641-52-3173	1/1/2024	4:59:00 AM	L. Harford	Female	37	White	NaN	Not Admission	5.0	10

- Initial Exploration: Used `df.info()` to check structure

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9216 entries, 0 to 9215
Data columns (total 11 columns):
 #   Column           Non-Null Count  Dtype  
 ---  --  
 0   Patient Id       9216 non-null   object  
 1   Patient Admission Date  9216 non-null   object  
 2   Patient Admission Time 9216 non-null   object  
 3   Merged            9216 non-null   object  
 4   Patient Gender    9216 non-null   object  
 5   Patient Age       9216 non-null   int64  
 6   Patient Race      9216 non-null   object  
 7   Department Referral 3816 non-null   object  
 8   Patient Admission Flag 9216 non-null   object  
 9   Patient Satisfaction Score 2517 non-null   float64 
 10  Patient Waittime   9216 non-null   int64  
dtypes: float64(1), int64(2), object(8)
memory usage: 792.1+ KB
```

❖ Used `df.duplicated()` to check duplicates

```
Patient Id          0
Patient Admission Date 0
Patient Admission Time 0
Merged              0
Patient Gender      0
Patient Age          0
Patient Race         0
Department Referral  5400
Patient Admission Flag 0
Patient Satisfaction Score 6699
Patient Waittime     0
dtype: int64
```

- **Date Time standardization**

Change column type of Patient Admission Date, Patient Admission Time to Date time column type

- **Remove Unwanted Columns**

Remove merged column

- **Standardization Column Value Names**

Correct Patient Gender to Female

- **Adding New Column**

Add new Column Age Group

- **Filling none Values**

Filling Department Referral by No referral

- **Change column values**

Change Patient Admission Flag to Yes / No

- **Filled Missing Values**

Filled missing values in Patient Satisfaction Score column using median

- **Change column name**

Change column name of Patient Wait time to Patient Wait time minutes

4. Data Analysis Using MYSQL

1. Average patient wait time across different age groups

	Age Group	avg_waittime
▶	Adult	35.25
	Senior	35.21
	Child	35.32

2. Dates were the highest patient admissions recorded

	Patient Admission Date	admissions
▶	12/4/2023	30
	21/05/2024	28
	1/8/2024	28
	15/12/2023	27
	7/3/2024	26
	22/06/2023	26
	15/09/2023	26
	1/6/2023	26
	18/10/2023	25
	17/08/2024	25
	14/01/2024	25
	12/7/2024	25
	30/05/2023	24
	3/3/2024	24
	2/3/2024	24
	10/8/2024	24
	10/8/2023	24
	10/5/2024	24
	5/8/2023	23
	3/6/2024	23
	28/06/2023	23
	22/09/2023	23
	20/08/2024	23
	15/06/2024	23

3. Patient satisfaction scores difference between genders

	Patient Gender	satisfaction_score
▶	Female	4.99
	Male	5.01

4. Department referrals lead to the longest average wait times

	Department Referral	wait_time
▶	neurology	36.80
	physiotherapy	36.57
	gastroenterology	35.83
	cardiology	35.35
	no referral	35.29
	orthopedics	34.98
	general practice	34.91
	renal	34.70

5. Proportion of admissions are flagged as Admission versus Not Admission

	Patient Admission Flag	admission_status
▶	No	4604
	Yes	4612

6. Hours of the day do patient admissions peaks

	hour	admissions
►	23	436
	7	415
	13	410
	0	406
	11	403
	15	394
	5	393
	9	388
	8	386
	3	385
	4	384
	19	383
	16	378
	2	376

7. Age group reports the highest average satisfaction scores

	Age Group	avg_score
►	Adult	5.02
	Child	4.99
	Senior	4.92

8. Do longer wait times correlate with lower patient satisfaction scores

	Patient Waittime mins	avg_score
►	10	4.99
	11	5.23
	12	5.25
	13	5
	14	5.07
	15	5.08
	16	4.95
	17	5.01
	18	5.12
	19	5.04
	20	4.9
	21	4.97

9. Patient race group has the highest admission frequency

	Patient Race	admission_frequency
►	White	2571
	African American	1951
	Two or More Races	1557
	Asian	1060
	Declined to Identify	1030
	Pacific Islander	549
	Native American/Alaska Native	498

10. department handles the largest patient volume overall

Department Referral	patient_volume
no referral	5400
general practice	1840
orthopedics	995
physiotherapy	276
cardiology	248
neurology	193
gastroenterology	178
renal	86

5. Data Visualization Using Power Bi

Build interactive dashboard in power bi to present insights



6. Recommendations

- Prioritize orthopedics and general Neurology, Physiotherapy practice with additional staff to handle **patient volume** and **reduce wait time**.
- Optimize gender-specific services for females **Leverage 47 %** Of female dominance with women-focused scheduling or amenities.
- Improve satisfaction for young and seniors: Target the smallest age group for **faster consultations**, as younger and senior patient satisfaction rate lower.
- Operational Enhancements
Introduce **mobile apps for real time updates** and **AI** in various areas to **optimize patient flow and reduce wait time**.