



Data Analysis of Health care system

Comprehensive analysis of patient visits, satisfaction, and department referrals to improve overall service quality.

Presented by :

Samaa Mohamed saeed
Mohamed Ashraf Esmail
Mahitab Yasser Yassin





We will go through four phases



1 Excel

using Tools and Techniques for Effective Data insights



2 Python

Use to clean and Transforming Data



3 Power Bi

.Data visualization Extracting useful insights

4 Report Visualization



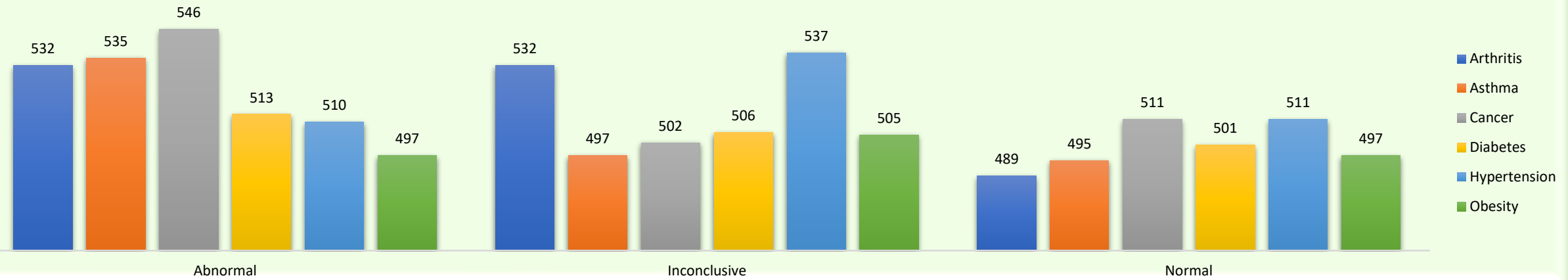
Sample of Data collection in Excel

| # | patient_id | patient_gender | patient_age | patient_sat_score | patient_first_initial | patient_last_name | patient_race | patient_admin_flag | patient_waittime | department_referral | Blood Type | Medication | Insurance Provider | Test Result | Medical Condition | Admission Type | Billing Amount | Doctor |
|----|-------------|----------------|-------------|-------------------|-----------------------|-------------------|------------------------------|--------------------|------------------|---------------------|------------|-------------|--------------------|--------------|-------------------|----------------|----------------|-------------------|
| 3 | 316-34-3057 | M | 4 | X | | Methuen | Native American/Alaska Nativ | TRUE | 27 | None | A+ | Ibuprofen | Medicare | Inconclusive | Obesity | Emergency | 33,643 | Samantha Davies |
| 4 | 897-46-3852 | F | 56 | 9 | P | Schubuser | African American | TRUE | 55 | General Practice | A- | Aspirin | Aetna | Normal | Obesity | Emergency | 27,955 | Tiffany Mitchell |
| 5 | 358-31-9711 | F | 24 | 8 | U | Titcombe | Native American/Alaska Nativ | TRUE | 31 | General Practice | O+ | Ibuprofen | Medicare | Abnormal | Diabetes | Elective | 37,910 | Kevin Wells |
| 6 | 289-26-0537 | M | 5 | Y | | Gionettitti | African American | FALSE | 10 | Orthopedics | AB+ | Penicillin | Aetna | Abnormal | Cancer | Urgent | 14,238 | Kathleen Hanna |
| 7 | 255-51-2877 | M | 58 | H | | Buff | Asian | FALSE | 59 | None | A+ | Ibuprofen | UnitedHealthcare | Normal | Asthma | Urgent | 48,145 | Taylor Newton |
| 8 | 465-97-0990 | F | 68 | F | | Perrat | White | TRUE | 43 | None | AB- | Paracetamol | Medicare | Inconclusive | Diabetes | Emergency | 19,581 | Kelly Olson |
| 9 | 157-31-7520 | F | 47 | K | | Gwillim | Two or More Races | TRUE | 23 | None | A+ | Paracetamol | Cigna | Inconclusive | Cancer | Emergency | 45,820 | Suzanne Thomas |
| 10 | 432-34-5614 | F | 79 | 1 | E | Dewhirst | White | FALSE | 42 | None | AB+ | Aspirin | Cigna | Abnormal | Asthma | Elective | 50,119 | Daniel Ferguson |
| 11 | 609-17-8678 | M | 62 | M | | Crebo | African American | FALSE | 51 | None | AB- | Paracetamol | UnitedHealthcare | Inconclusive | Cancer | Elective | 19,785 | Heather Day |
| 12 | 497-14-6812 | F | 73 | Q | | Churchard | White | TRUE | 34 | Gastroenterology | O+ | Paracetamol | Medicare | Normal | Cancer | Urgent | 12,577 | John Duncan |
| 13 | 393-38-9502 | F | 16 | N | | Corpes | White | FALSE | 39 | Orthopedics | A- | Lipitor | Medicare | Inconclusive | Hypertension | Urgent | 8,000 | Douglas Mayo |
| 14 | 288-05-6370 | F | 16 | R | | Brixey | Native American/Alaska Nativ | TRUE | 53 | General Practice | A+ | Penicillin | Cigna | Abnormal | Diabetes | Emergency | 43,282 | Kenneth Fletcher |
| 15 | 784-54-9931 | M | 46 | M | | Goudie | Pacific Islander | FALSE | 45 | None | AB+ | Lipitor | UnitedHealthcare | Normal | Asthma | Urgent | 33,208 | Theresa Freeman |
| 16 | 662-21-6522 | M | 69 | G | | Stanlack | White | TRUE | 49 | None | AB+ | Paracetamol | UnitedHealthcare | Normal | Cancer | Urgent | 40,702 | Roberta Stewart |
| 17 | 628-73-1801 | M | 37 | C | | McMurty | Declined to Identify | TRUE | 57 | None | AB- | Aspirin | Blue Cross | Inconclusive | Obesity | Elective | 12,263 | Maria Dougherty |
| 18 | 370-19-2271 | F | 50 | I | | Scothorn | Asian | FALSE | 35 | General Practice | A+ | Paracetamol | UnitedHealthcare | Normal | Arthritis | Elective | 24,500 | Erica Spencer |
| 19 | 458-98-8860 | M | 37 | J | | Helgass | Declined to Identify | FALSE | 55 | None | AB- | Paracetamol | Cigna | Abnormal | Obesity | Urgent | 17,440 | Justin Kim |
| 20 | 728-31-2493 | F | 70 | W | | Chittock | Asian | TRUE | 50 | Physiotherapy | A- | Aspirin | Blue Cross | Abnormal | Hypertension | Elective | 18,843 | Justin Moore Jr. |
| 21 | 823-34-5523 | M | 55 | 2 | F | Prendergast | Asian | TRUE | 40 | None | B+ | Penicillin | Cigna | Normal | Asthma | Elective | 23,762 | Denise Galloway |
| 22 | 621-70-7472 | F | 63 | T | | Bissiker | Native American/Alaska Nativ | TRUE | 25 | None | A- | Aspirin | Blue Cross | Normal | Asthma | Elective | 43 | Krista Smith |
| 23 | 344-36-7156 | F | 44 | 2 | M | Mandell | Asian | FALSE | 51 | None | B+ | Lipitor | Aetna | Normal | Asthma | Urgent | 17,696 | Gregory Smith |
| 24 | 455-21-3671 | F | 11 | 4 | D | Coste | Declined to Identify | FALSE | 30 | None | O+ | Lipitor | Aetna | Inconclusive | Arthritis | Urgent | 5,998 | Vanessa Newton |
| 25 | 259-10-0339 | M | 8 | Q | | Dodridge | Asian | FALSE | 16 | General Practice | A+ | Lipitor | Aetna | Inconclusive | Asthma | Urgent | 25,250 | Donna Martinez M |
| 26 | 720-54-2625 | F | 4 | C | | Pavie | Native American/Alaska Nativ | TRUE | 23 | None | AB+ | Paracetamol | Medicare | Inconclusive | Hypertension | Emergency | 33,211 | Stephanie Kramer |
| 27 | 661-92-7059 | M | 42 | 0 | Z | Sleightholm | African American | FALSE | 51 | None | AB+ | Aspirin | Medicare | Abnormal | Obesity | Urgent | 19,747 | Angela Contreras |
| 28 | 598-53-3927 | F | 68 | S | | Noads | African American | FALSE | 58 | None | O- | Ibuprofen | Blue Cross | Normal | Asthma | Elective | 26,787 | Wendy Glenn |
| 29 | 715-74-5338 | M | 22 | J | | Filkin | Pacific Islander | FALSE | 25 | General Practice | A- | Lipitor | UnitedHealthcare | Abnormal | Diabetes | Emergency | 18,835 | James Ellis |
| 30 | 669-74-2146 | F | 58 | C | | Bilby | Declined to Identify | FALSE | 55 | Orthopedics | O- | Penicillin | Cigna | Normal | Hypertension | Emergency | 32,643 | Jared Bruce Jr. |
| 31 | 693-38-2084 | F | 72 | H | | Kehoe | African American | TRUE | 37 | None | B+ | Paracetamol | Aetna | Abnormal | Hypertension | Elective | 5,767 | Brandy Mitchell |
| 32 | 548-93-9953 | F | 73 | S | | O'Neill | Two or More Races | FALSE | 46 | Neurology | B- | Ibuprofen | Cigna | Inconclusive | Asthma | Urgent | 47,909 | Jacob Huynh |
| 33 | 334-76-4005 | M | 35 | J | | Sibbit | White | TRUE | 40 | Orthopedics | A+ | Penicillin | UnitedHealthcare | Abnormal | Arthritis | Urgent | 25,835 | Kristina Frazier |
| 34 | 846-66-7490 | M | 48 | T | | Clissett | White | FALSE | 53 | None | AB- | Penicillin | Medicare | Normal | Arthritis | Elective | 17,993 | John Hartman |
| 35 | 278-49-6531 | F | 74 | 0 | L | Vannacci | White | FALSE | 20 | Physiotherapy | AB- | Ibuprofen | Medicare | Normal | Cancer | Elective | 21,186 | Heather Garcia |
| 36 | 600-62-3412 | M | 14 | H | | Ponting | Two or More Races | TRUE | 52 | None | AB+ | Lipitor | Blue Cross | Normal | Diabetes | Emergency | 8,409 | Lynn Young |
| 37 | 745-62-9973 | F | 37 | D | | Bleacher | Declined to Identify | TRUE | 49 | None | AB+ | Lipitor | Medicare | Abnormal | Asthma | Elective | 45,453 | Emma Allison |
| 38 | 384-70-5840 | M | 25 | B | | Drinkeld | African American | FALSE | 10 | None | AB+ | Paracetamol | Aetna | Inconclusive | Diabetes | Urgent | 4,398 | Laura Myers |
| 39 | 285-41-0221 | M | 39 | N | | Aggiss | Two or More Races | FALSE | 52 | Orthopedics | B- | Ibuprofen | Aetna | Normal | Hypertension | Urgent | 9,380 | Travis Parsons |
| 40 | 124-73-3676 | M | 48 | K | | Bilofsky | Declined to Identify | FALSE | 35 | None | O- | Lipitor | Cigna | Abnormal | Diabetes | Emergency | 48,291 | Christine Johnson |
| 41 | 878-55-0051 | F | 22 | R | | Harrington | African American | FALSE | 16 | General Practice | B- | Penicillin | Medicare | Abnormal | Obesity | Urgent | 32,974 | Emily Taylor |
| 42 | 206-25-4281 | F | 72 | 7 | D | Yarrington | Declined to Identify | TRUE | 33 | None | B- | Lipitor | Cigna | Normal | Hypertension | Emergency | 49,943 | William Reynolds |
| 43 | 472-28-5842 | F | 41 | N | | Gulliman | Pacific Islander | FALSE | 16 | None | A+ | Aspirin | Cigna | Normal | Hypertension | Urgent | 35,634 | Jennifer Carter |
| 44 | 553-27-8559 | F | 8 | Q | | Garnsworthy | African American | FALSE | 59 | General Practice | A- | Paracetamol | Cigna | Normal | Asthma | Elective | 25,966 | Matthew Thomas |
| 45 | 319-83-9008 | F | 70 | H | | Hesbrook | Native American/Alaska Nativ | TRUE | 39 | None | AB+ | Paracetamol | UnitedHealthcare | Normal | Obesity | Urgent | 21,784 | Mary Logan |
| 46 | 434-16-5502 | F | 50 | N | | Romanet | White | TRUE | 37 | None | AB- | Penicillin | Medicare | Abnormal | Cancer | Urgent | 42,685 | Kristin Martinez |
| 47 | 672-89-2190 | F | 33 | 3 | K | Rushmare | Declined to Identify | TRUE | 16 | None | O+ | Penicillin | Medicare | Normal | Cancer | Emergency | 45,586 | Daniel Murphy |
| 48 | 135-05-4896 | M | 17 | P | | Pow | Two or More Races | FALSE | 56 | None | O+ | Paracetamol | Cigna | Normal | Hypertension | Elective | 29,615 | Emily Patterson |
| 49 | 760-33-4176 | F | 45 | Q | | Dear | White | FALSE | 57 | Orthopedics | A- | Aspirin | Cigna | Normal | Diabetes | Urgent | 36,992 | Heather Smith |
| 50 | 624-71-9493 | F | 53 | X | | Leake | White | FALSE | 21 | None | O+ | Ibuprofen | UnitedHealthcare | Abnormal | Hypertension | Urgent | 28,051 | John Smith |

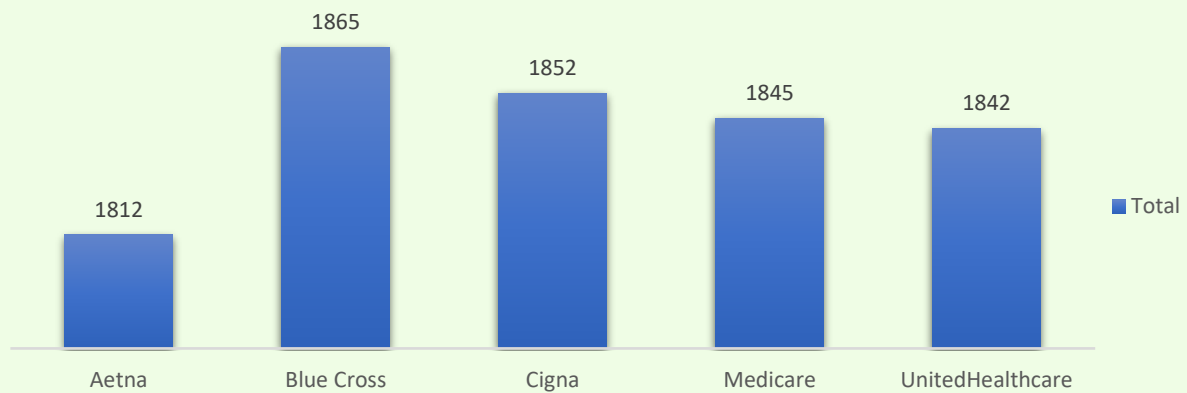


Sample of Dash board in Excel

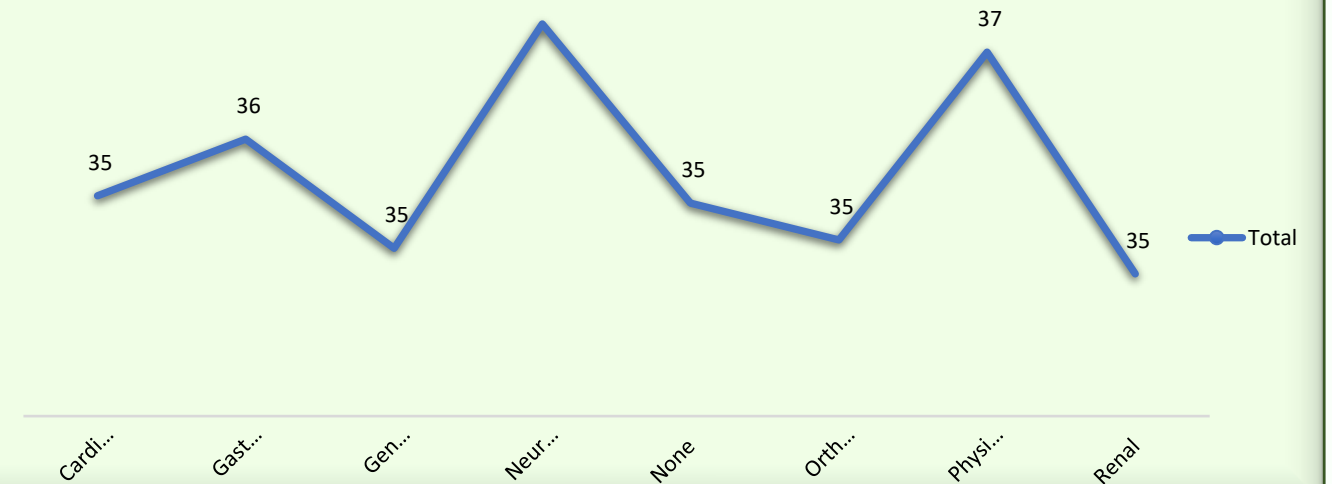
count of patient id of Health Patterns Associated with Medical Conditions Based on Test Results



count patient id of each insurance provider

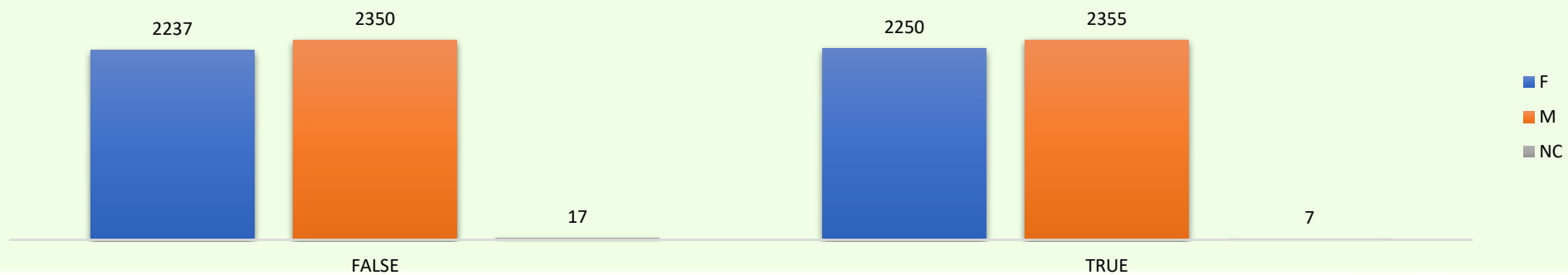


Average of patient waittime of each department_referral

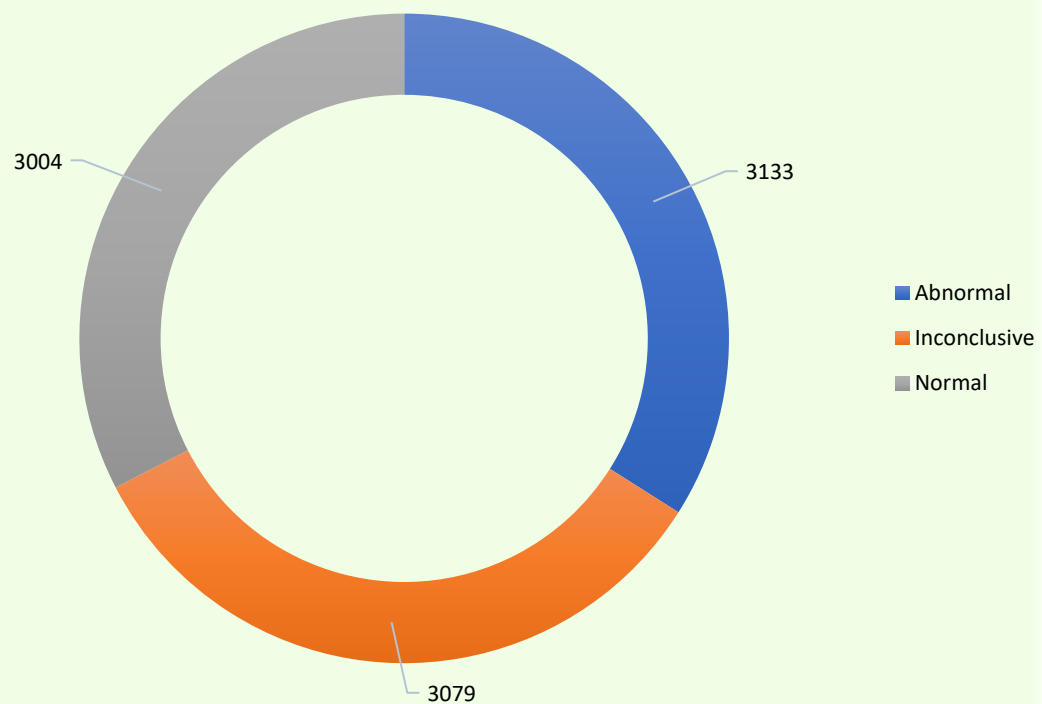




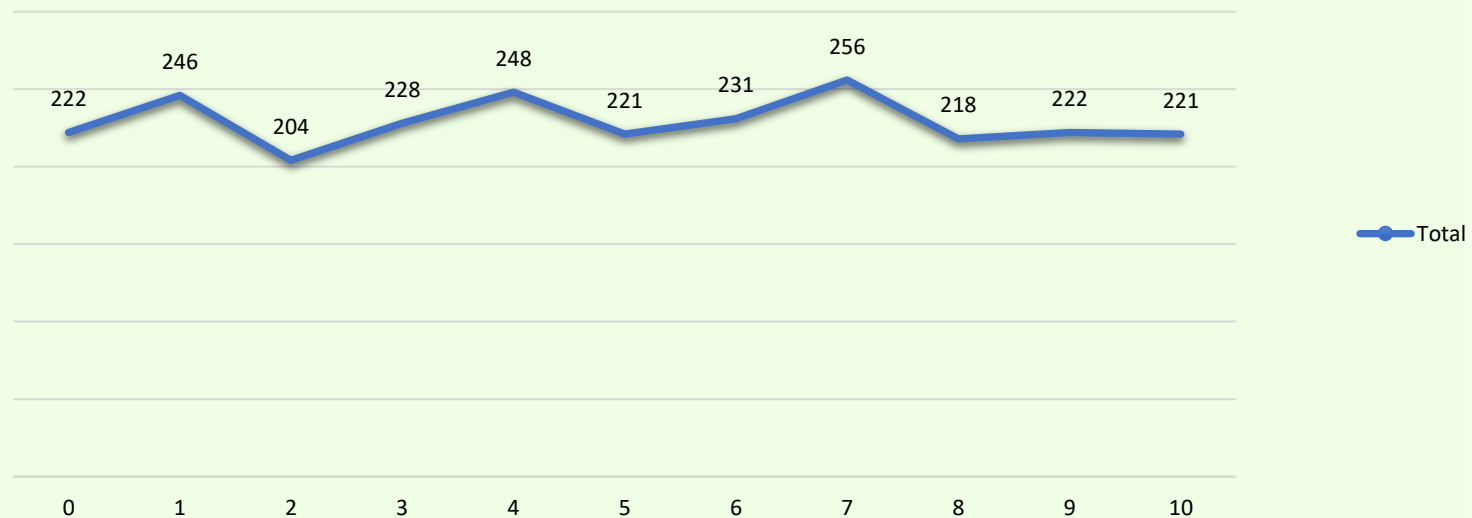
count of patient that have admin_flag according gender



count of patient of each test results



number of patient of patient_satisfaction score





Data cleaning and Transforming by python

```
# Import necessary libraries
import pandas as pd

# Load the dataset (replace 'your_file.csv' with the actual path)
df = pd.read_csv('C:/Users/DELL/Desktop/depi.csv')

# Display the first few rows of the dataset
df.head()
```

Python

| | date | patient_id | patient_gender | patient_age | patient_sat_score | patient_first_initial | patient_last_name | patient_race | patient_admin_flag | patient_waittime | department_referral |
|---|---------------------|-----------------|----------------|-------------|-------------------|-----------------------|-------------------|----------------------------------|--------------------|------------------|---------------------|
| 0 | 20/03/2020 8:47 | 145-39- 5406 | M | 69 | 10.0 | H | Glasspool | White | False | 39 | NaN |
| 1 | 15/06/2020 11:29 | 316-34- 3057 | M | 4 | NaN | X | Methuen | Native American/Alaska Native | True | 27 | NaN |
| 2 | 20/06/2020 9:13 | 897-46- 3852 | F | 56 | 9.0 | P | Schubuser | African American | True | 55 | General Practice |
| 3 | 04/02/2020 22:34 | 358-31- 9711 | F | 24 | 8.0 | U | Titcombe | Native American/Alaska Native | True | 31 | General Practice |
| 4 | 04/09/2020 17:48 | 289-26- 0537 | M | 5 | NaN | Y | Gionettitti | African American | False | 10 | Orthopedics |

```
# Convert date column to datetime format
df['date'] = pd.to_datetime(df['date'], format='%d/%m/%Y %H:%M')

# Convert 'patient_admin_flag' from string to boolean
df['patient_admin_flag'] = df['patient_admin_flag'].apply(lambda x: True if x == 'TRUE' else False)

# Replace missing SAT scores with 0 (if it's reasonable, otherwise impute differently)
df['patient_sat_score'].fillna(0, inplace=True)

# Check for any missing data
print(df.isnull().sum())

# Drop columns that are not useful (like patient_first_initial, patient_last_name)
df = df.drop(['patient_first_initial', 'patient_last_name'], axis=1)
```

Python

```
date                0
patient_id          0
patient_gender      0
patient_age         0
patient_sat_score   0
patient_first_initial 0
patient_last_name   0
patient_race        0
patient_admin_flag  0
patient_waittime    0
department_referral 5400
dtype: int64
```

```

from sklearn.preprocessing import LabelEncoder

# Encode categorical features
label_encoder = LabelEncoder()
df['patient_gender'] = label_encoder.fit_transform(df['patient_gender'])
df['patient_race'] = label_encoder.fit_transform(df['patient_race'])
df['department_referral'] = label_encoder.fit_transform(df['department_referral'])

# Check the transformed data
df.head()

```

| | date | patient_id | patient_gender | patient_age | patient_sat_score | patient_race | patient_admin_flag | patient_waittime | department_referral |
|---|---------------------|-------------|----------------|-------------|-------------------|--------------|--------------------|------------------|---------------------|
| 0 | 2020-03-20 08:47:00 | 145-39-5406 | 1 | 69 | 10.0 | 6 | False | 39 | 7 |
| 1 | 2020-06-15 11:29:00 | 316-34-3057 | 1 | 4 | 0.0 | 3 | False | 27 | 7 |
| 2 | 2020-06-20 09:13:00 | 897-46-3852 | 0 | 56 | 9.0 | 0 | False | 55 | 2 |
| 3 | 2020-02-04 22:34:00 | 358-31-9711 | 0 | 24 | 8.0 | 3 | False | 31 | 2 |
| 4 | 2020-09-04 17:48:00 | 289-26-0537 | 1 | 5 | 0.0 | 0 | False | 10 | 4 |

```

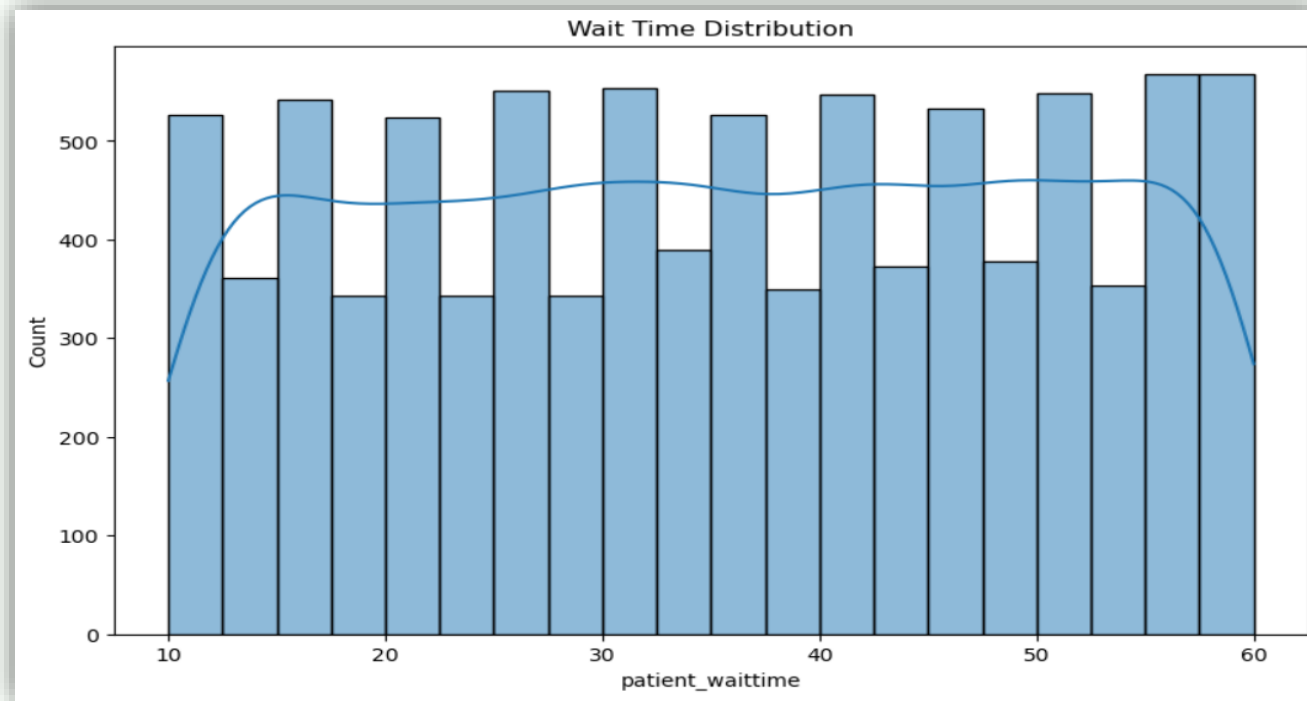
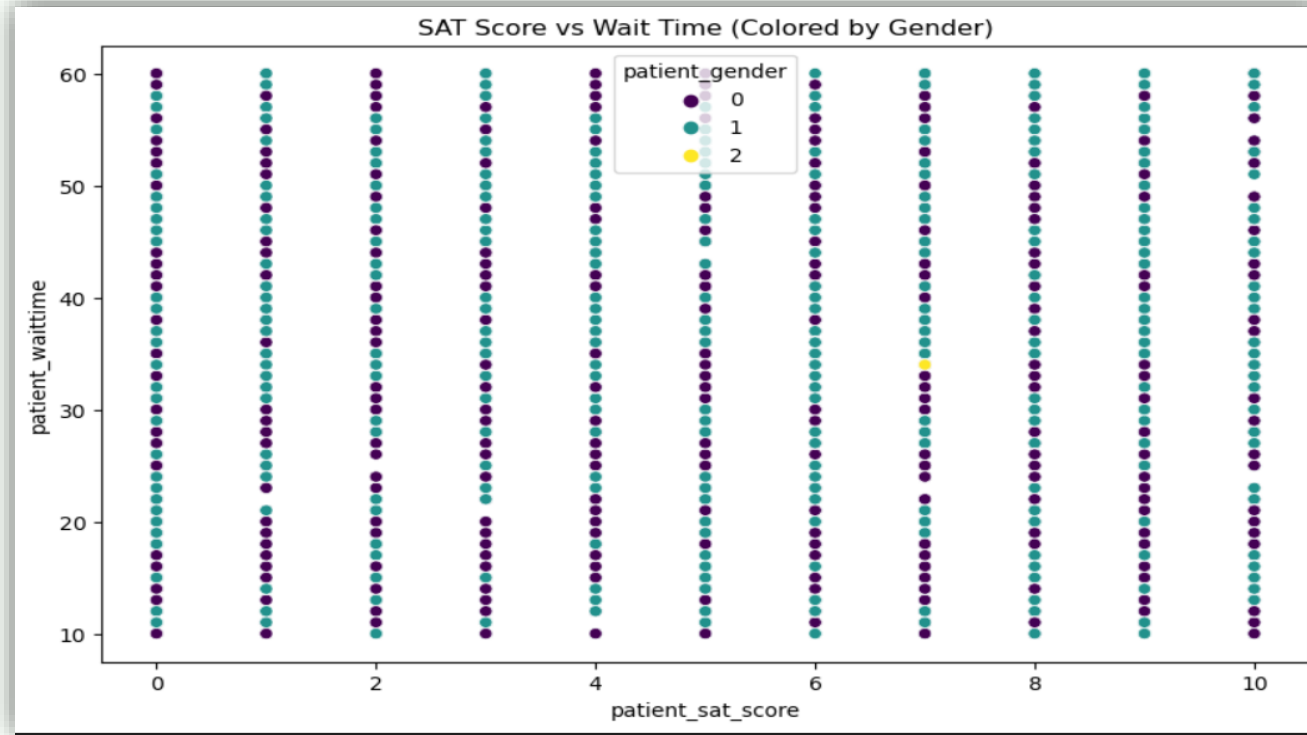
import matplotlib.pyplot as plt
import seaborn as sns

# Plot the distribution of patient ages
plt.figure(figsize=(10, 6))
sns.histplot(df['patient_age'], bins=20, kde=True)
plt.title('Age Distribution of Patients')
plt.show()

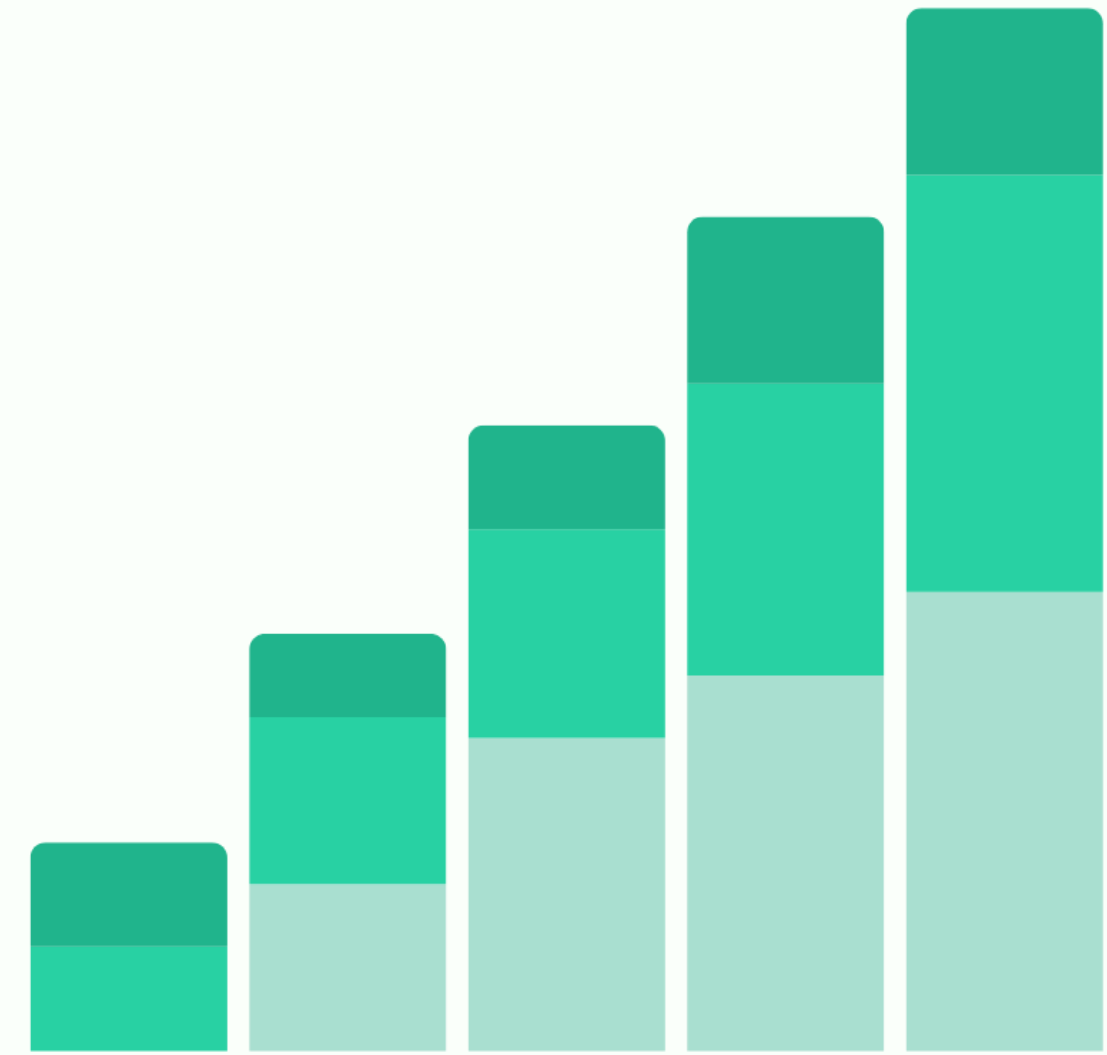
# Plot the distribution of patient wait times
plt.figure(figsize=(10, 6))
sns.histplot(df['patient_waittime'], bins=20, kde=True)
plt.title('Wait Time Distribution')
plt.show()

# Visualize relationship between SAT score and wait time
plt.figure(figsize=(10, 6))
sns.scatterplot(data=df, x='patient_sat_score', y='patient_waittime', hue='patient_gender', palette='viridis')
plt.title('SAT Score vs Wait Time (Colored by Gender)')
plt.show()

```

Sample of data that analyzed by Python





Data Visualization by Power Bi



Patient Emergency Room Visit Report

Total Patients Visit

9,216

50.04%

Administrative Appointment

49.96%

None-Administrative



Avg. Satisfaction
5.47



Service Not Rated
75.10%



Avg. Wait Time
35.26



Referred Patients
41.41%



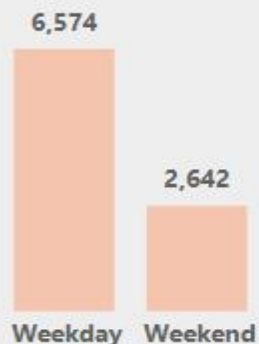
Walk In Patients
58.59%

Select The Moment

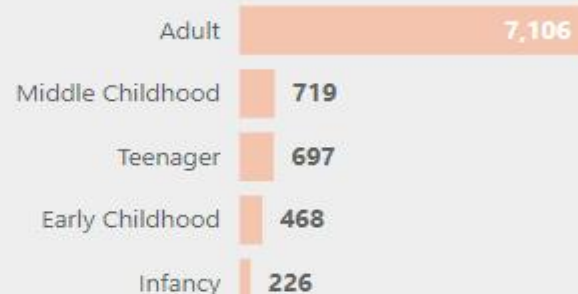
AM

PM

Patients by WeekType



Total Patients by Age Group



Total Patients Visits By Year



Total Patients by department_referral



Total Patients Visits



51.1%
Male



48.69%
Female

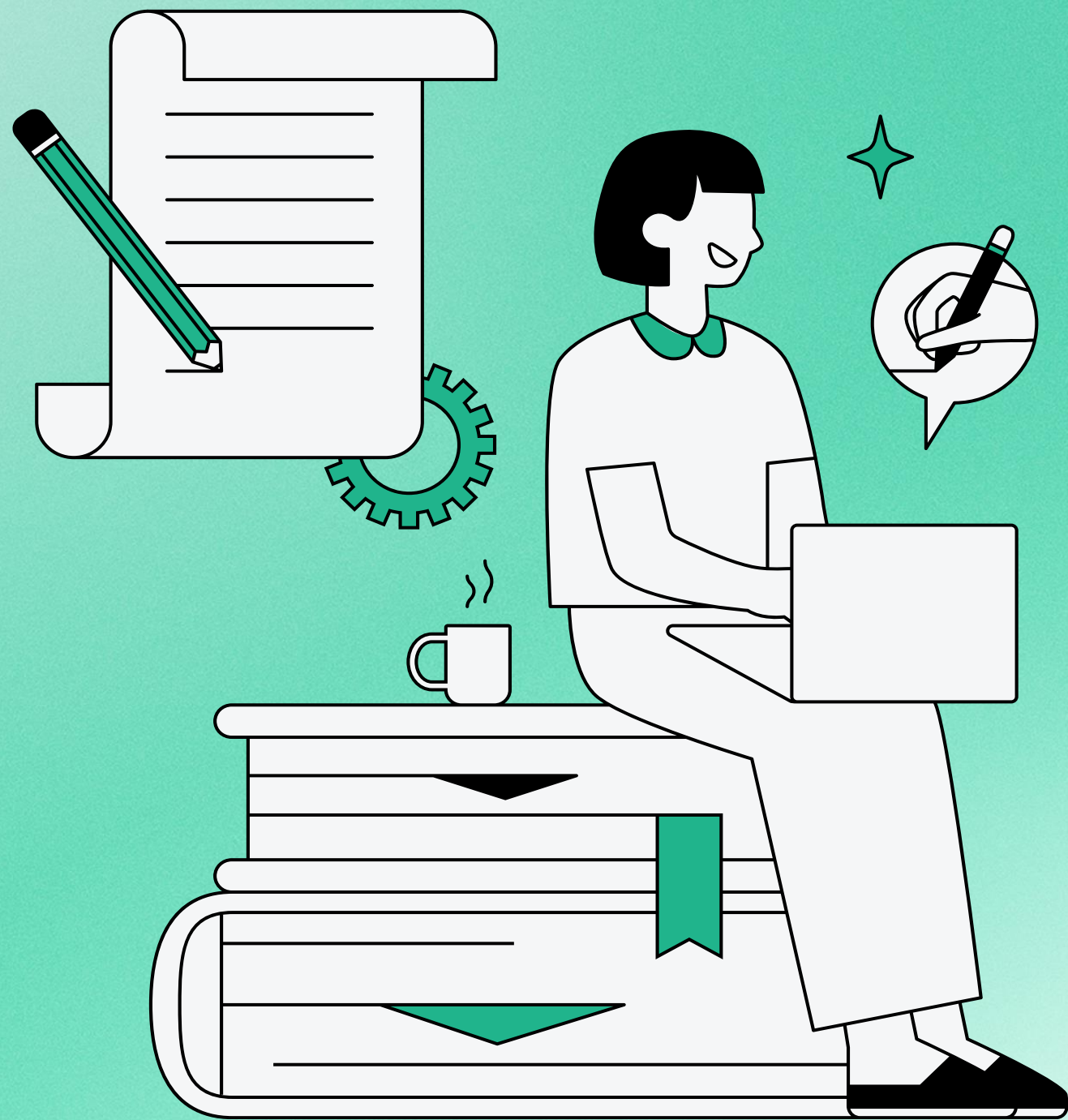


0.3%
Other

Patients are most SATISFIED when the SCALE shows the darkest GREEN on the Age-Group

Avg. Wait Time

| patient_race | 0-10 | 11-20 | 21-30 | 31-40 | 41-50 | 51-60 | 61-70 | 70+ |
|-------------------------------|------|-------|-------|-------|-------|-------|-------|-----|
| African American | | | | | | | | |
| Asian | | | | | | | | |
| Declined to Identify | | | | | | | | |
| Native American/Alaska Native | | | | | | | | |
| Pacific Islander | | | | | | | | |
| Two or More Races | | | | | | | | |
| White | | | | | | | | |



Summery of Final Patient Emergency Room Visit Report

Patient Volume and Trends

Total Visits

9,216 patients, with 50.04% administrative and 49.96% non-administrative.

Weekday vs Weekend

6,574 weekday visits, 2,642 weekend visits.

Demographics

7,106 adult visits, 226 infant visits, 51.1% male, 48.69% female.



Patient Emergency Room Visit Report

Total Patients Visit

9,216

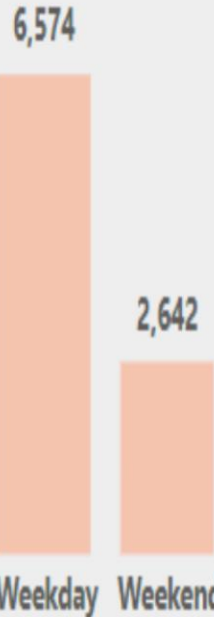
50.04%

Administrative
Appointment

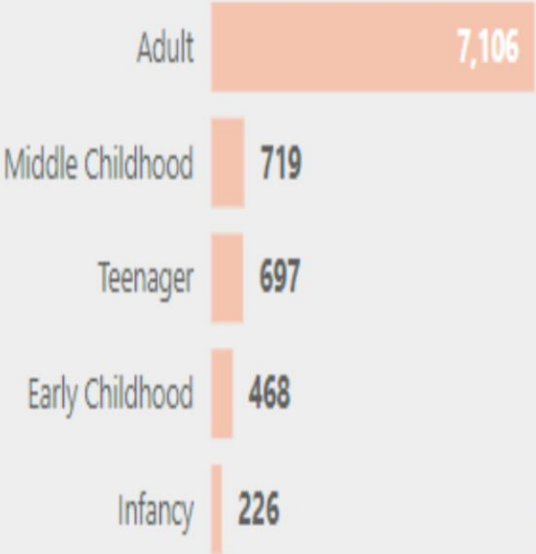
49.96%

None -
Administrative

Patients by WeekType



Total Patients by Age Group



Satisfaction and Wait Times



1

Satisfaction

Average of 5.47, with 75.10% not providing feedback.

2

Wait Times

Average of 35.26 minutes, with variations across age and race.

3

Opportunities

Increase feedback collection and target satisfaction initiatives.



Avg.
Satisfaction
5.47



Service Not
Rated
75.10%



Avg. Wait
Time
35.26

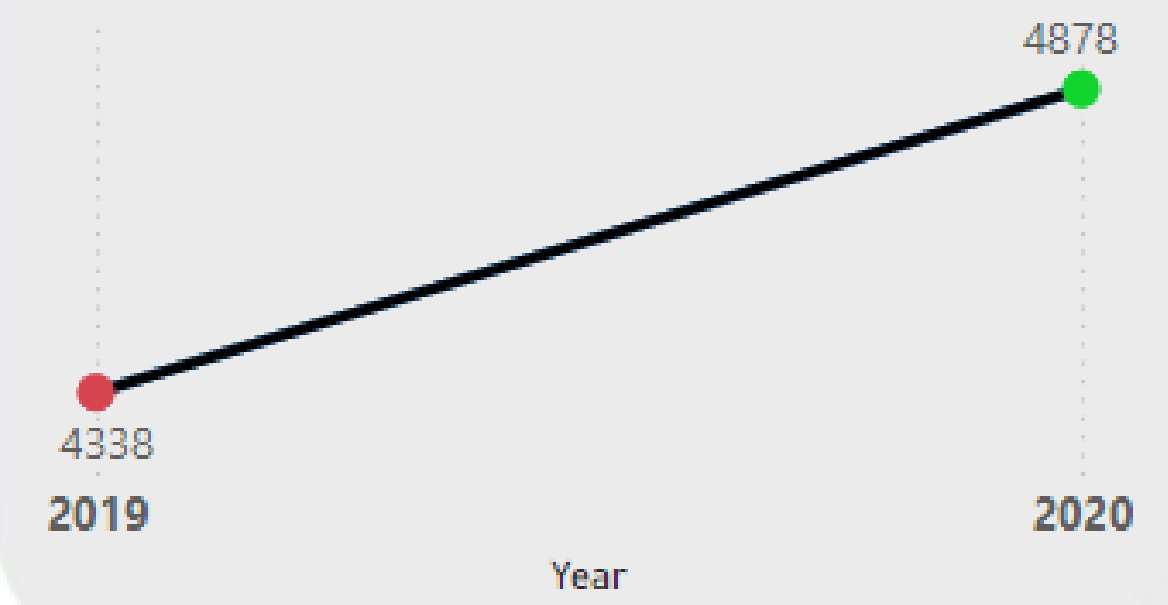




Resource Planning for Growth

| | | |
|----------------|---|-------------|
| Patient Visits | 2019: 4,338 | 2020: 4,878 |
| Increase | 12.4% year-over-year growth | |
| Opportunities | Scale resources to meet growing demand, analyze peak periods. | |

Total Patients Visits By Year





Thank
you very
much!

