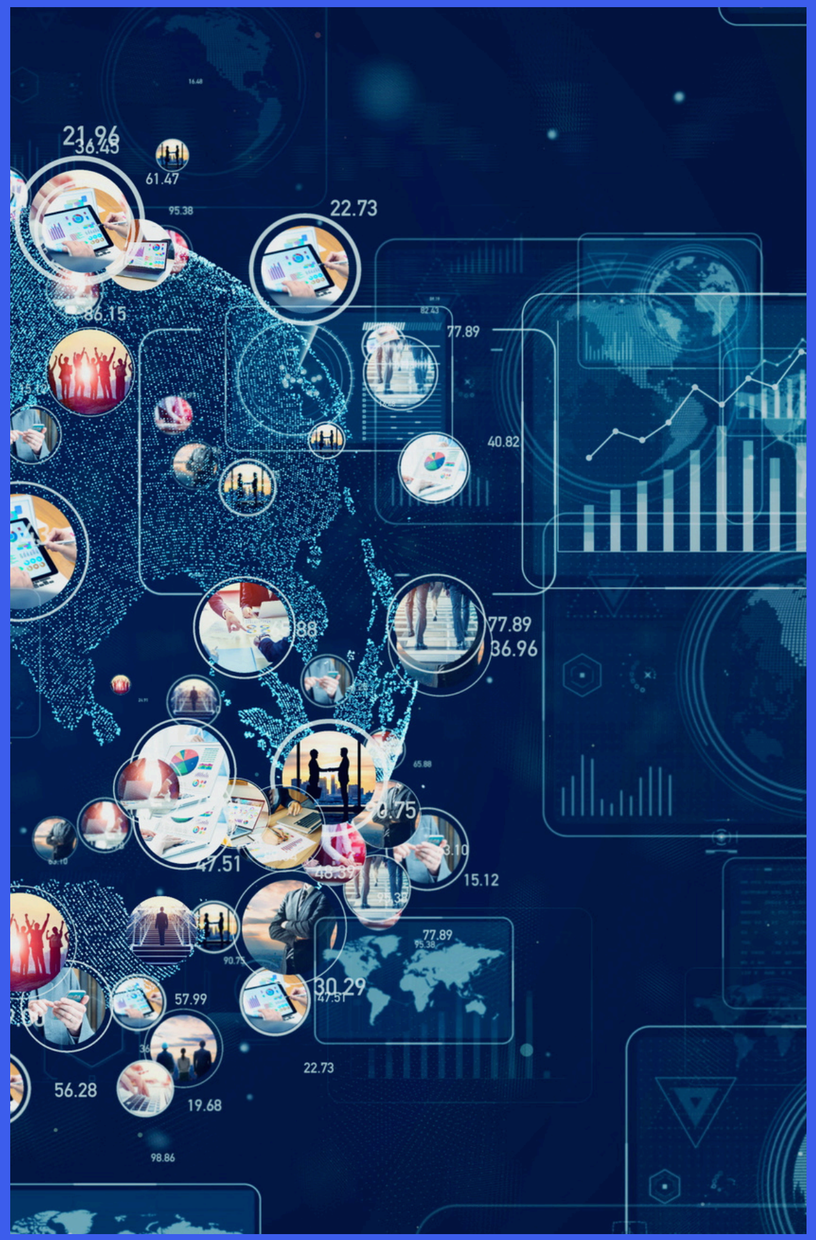


PHC4
Transparent Health Care Reporting

Data Literacy

A learning module from the
Pennsylvania Health Care Cost Containment Council

Jake Muskovitz
Special Projects

Kathrine Jordan
Communications & Education

September 26, 2025

About PHC4

PHC4 is governed by Council members who help oversee and carry out its mission. The Council is comprised of a multidisciplinary team of professionals that represent key stakeholders from across the Commonwealth.

Created in 1986 by the Pennsylvania General Assembly

Committees and Groups

- Technical Advisory Group
- Payment Data Advisory Group
- Data Systems Committee
- Education Committee
- Mandated Benefits Review Committee

PHC4

Transparent Health Care Reporting

PHC4's Data

Discover what PHC4's data can do for you

Review PHC4's Available Data, Data Price Structure, & identify your needs

Submit a Data Requests Application

Application review & approval process

Upon approval, analyst produces requested files. Client is provided with cost quote & invoice.

When payment is received, data files are securely transferred to client.
It's that easy!

PHC4 collects approximately 1.5 million inpatient records from Pennsylvania hospitals and approximately 3.5 million ambulatory/outpatient records from Pennsylvania hospitals and ambulatory surgery centers each year. These records are available for those who wish to analyze detailed inpatient discharge and ambulatory/outpatient procedure data in the form of standard predetermined data files or custom data files. There are more than 70 data fields available, which include utilization and administrative data.

PHC4
Transparent Health Care Reporting

Understanding Data

PHC4
Transparent Health Care Reporting

- Data literacy covers the fundamentals of understanding information
- Data fluency is data literacy; one's ability to read, write, and comprehend data effectively, performed at scale
- Data fluency includes familiarity with statistical methodologies and designs, including their limitations, and how they inform evidence-based actions

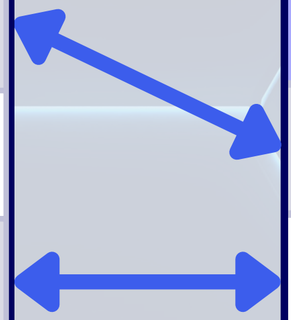
The purpose of this module is to move through data literacy toward data fluency!

A Foundation of Fluency

PHC4
Transparent Health Care Reporting

- All data is stored and analyzed in tables
  - Tables are made up of rows, sometimes called records, and columns, sometimes called variables
  - Databases commonly contain multiple tables that relate to each other in specified ways

Physician	Retired?	License Type
Dr. John Doe	FALSE	MD
Dr. Joe Smith	TRUE	MD
Dr. Jane Doe	FALSE	DO



Physician	Patient Census 2022	Patient Census 2023	Patient Census 2024
Dr. John Doe	400	421	408
Dr. Jane Doe	362	387	434

A Foundation of Fluency

PHC4
Transparent Health Care Reporting

- **Tables can be organized in long or wide format**
- **Cleaning, manipulating, analyzing, and visualizing data is commonly performed in software like Excel, Python, or R**

Physician	Patient Census 2022	Patient Census 2023	Patient Census 2024
Dr. John Doe	400	421	408
Dr. Jane Doe	362	387	434

Physician	Year	Patient Census
Dr. John Doe	2022	400
Dr. John Doe	2023	421
Dr. John Doe	2024	408
Dr. Jane Doe	2022	362
Dr. Jane Doe	2023	387
Dr. Jane Doe	2024	434

A Foundation of Fluency

Type of Data	Example	Note
Numeric	42 or 4.2	Numbers can be stored as integers or with decimals
Text	“PHC4”	Some databases impose limits on the number of characters for storage or organizational reasons
Boolean	True or False	
Datetime	“2025-04-15 17:30:00 EST”	Date and time are also commonly stored separately
Geographic	“POINT (-79.xxx, 40.xxx)“	Geographic data commonly has its own format including specific points and entire areas. Geographic data often requires specialized software to interpret

A Foundation of Fluency

PHC4
Transparent Health Care Reporting

Identifying Trustworthy Data

Structuring the Search

- **Population -** Who?
- **Intervention -** What or how?
- **Comparison -** Compared to what?
- **Outcome** - What are you trying to accomplish/improve/change?
- **Context** - In what kind of circumstances?

Considering the Sources

- Is this source trustworthy?
- Is there documentation explaining the format and collection method?
- Has the data been processed or is it raw?

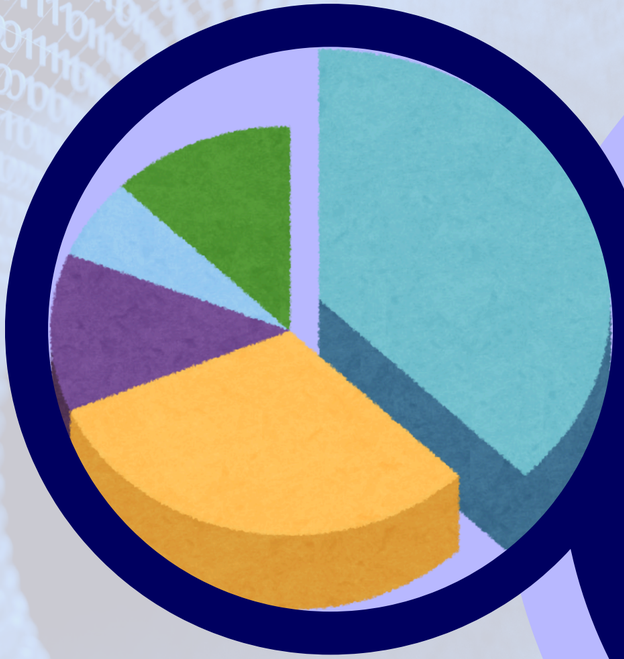
Discussion Question

Can you think of an example of trustworthy data and an example of untrustworthy data that you've come across?

What made it trustworthy or untrustworthy?

PHC4
Transparent Health Care Reporting

Interpreting Data

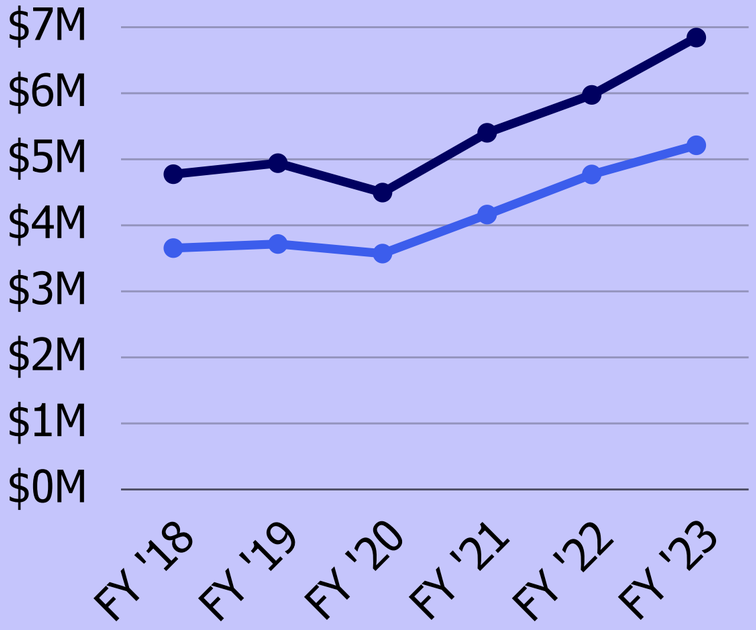


Understanding data visualizations, like graphs and charts, along with identifying patterns and drawing meaningful conclusions, are all key components in data fluency and communication.

Key

- Net Patient Revenue (NPR)
- Total Operating Expenses (TOE)

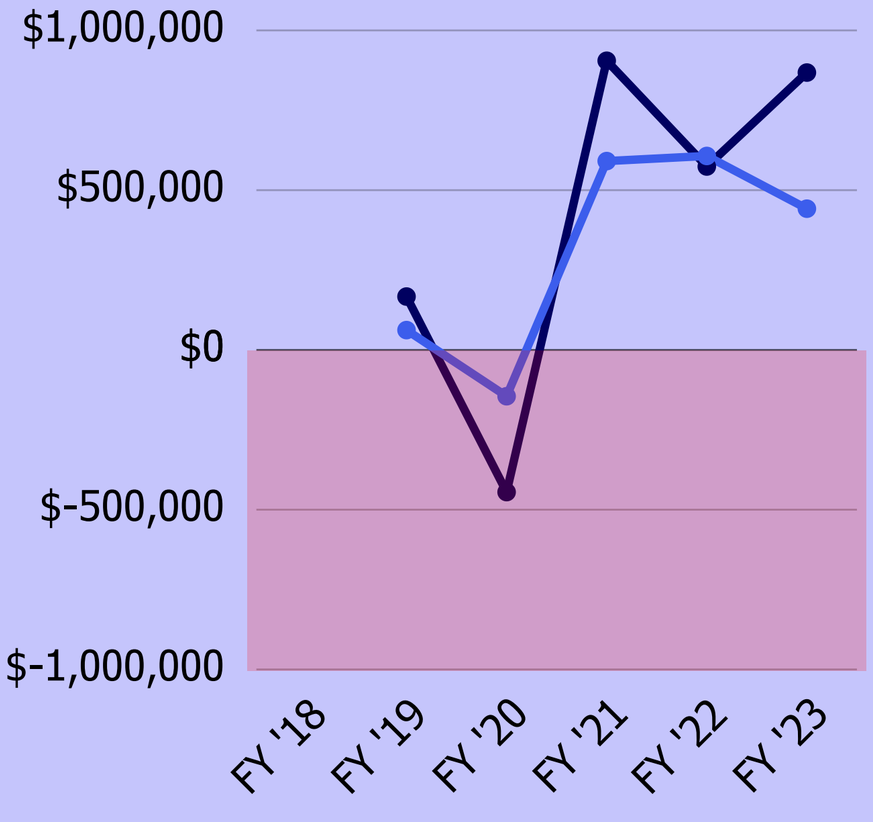
Statewide Average



Calculation:

X_i

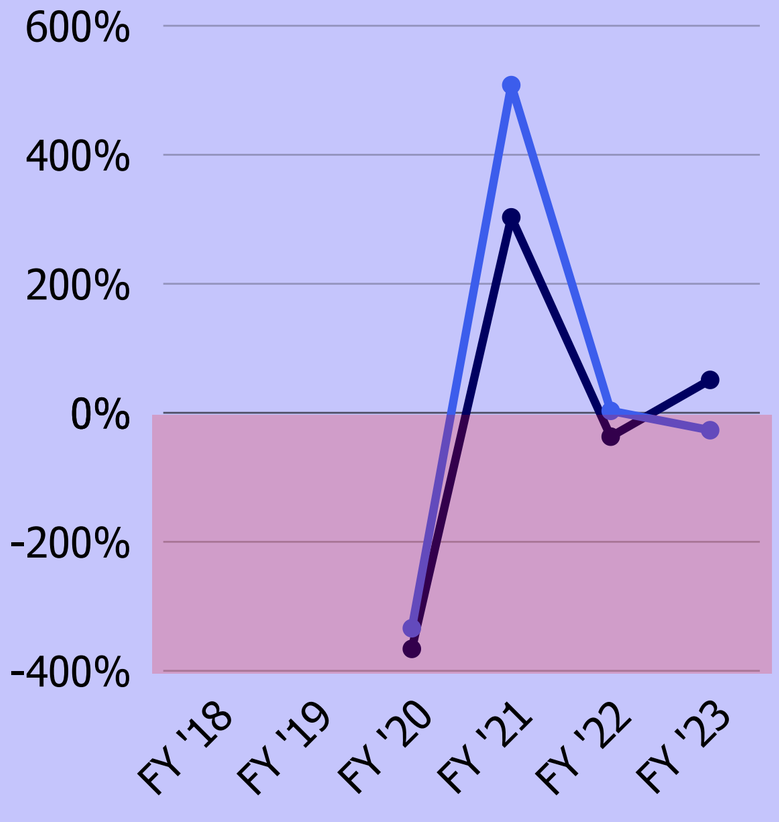
Annual Net Change



Calculation:

$Y_i = X_i - X_{i-1}$

Percent Change vs. Prior Year



Calculation:

$Z_i = (Y_i - Y_{i-1}) / Y_{i-1}$

Health Care Organization Dashboard

Registered Patients

.02%

247

Completed Appointments

13%

78

Referrals

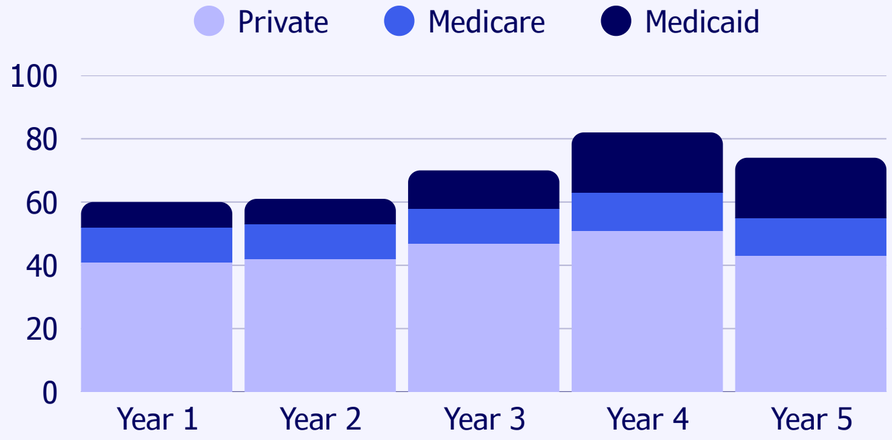
14%

50

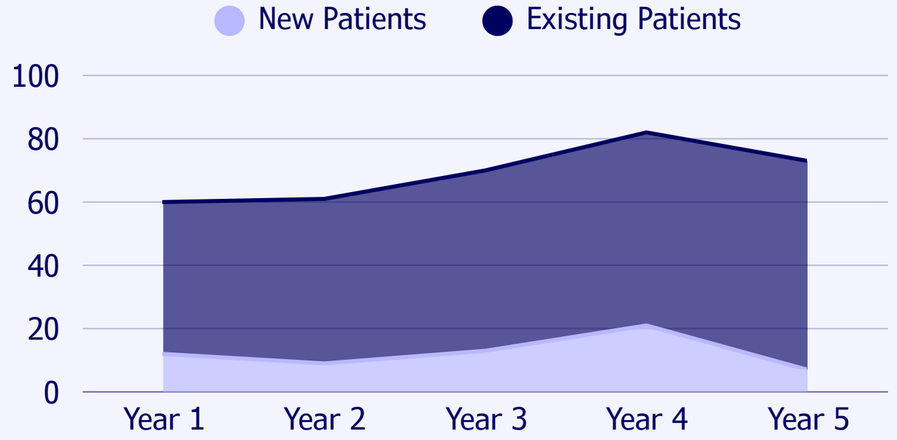
Online Consults

56%

32



Total Patient Population by Payor

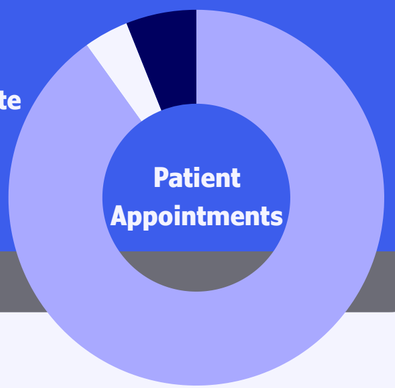


Total Patient Population Over the Most Recent 5 Years

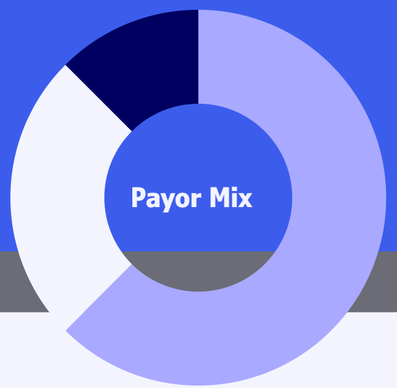
56% Total

56% Total

90.1%
Completion Rate



56% Total



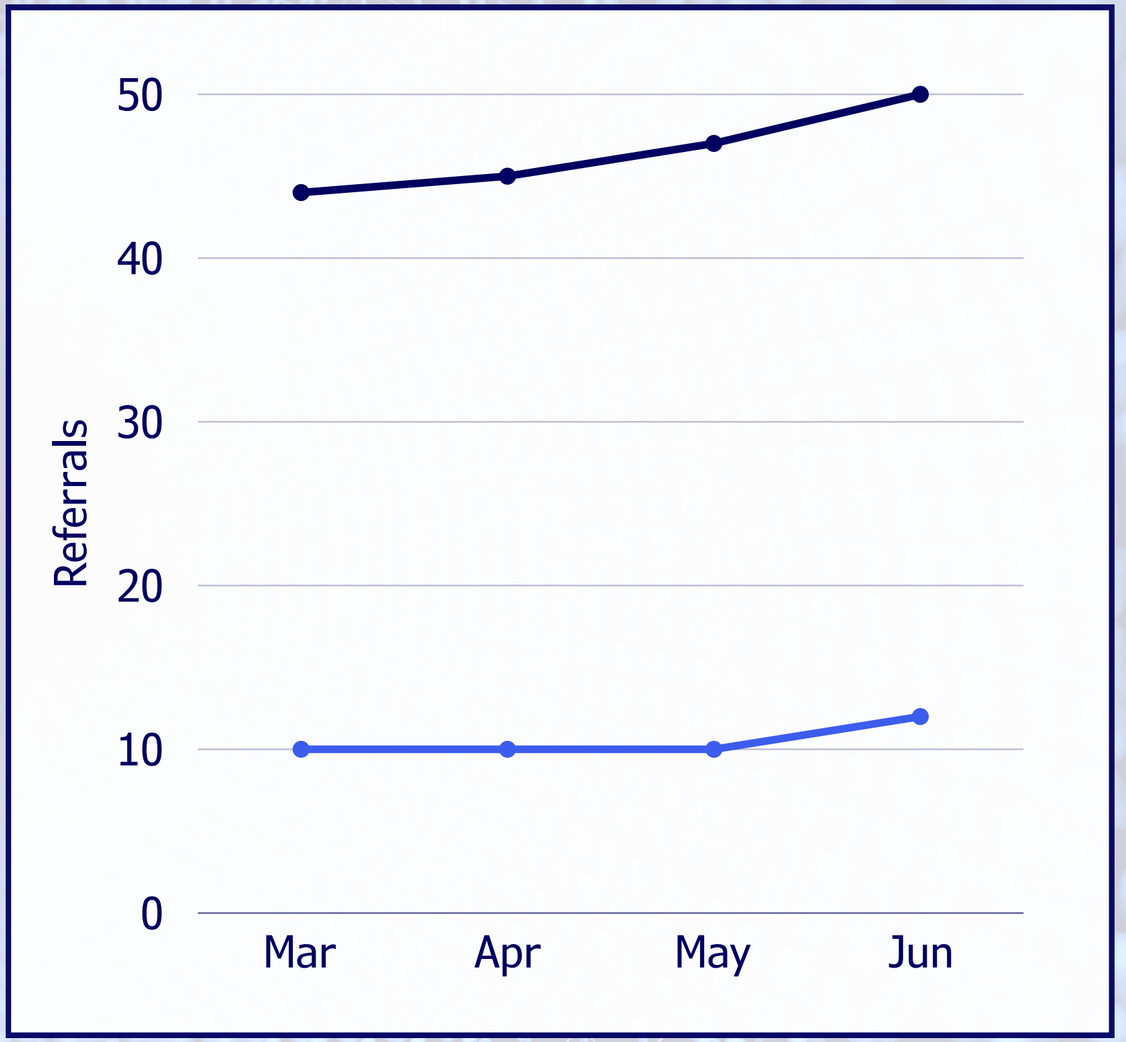
PHC4
Transparent Health Care Reporting

Referrals

14%

50ytd

**growth since March calculated monthly*



Health Care Organization Dashboard

Registered Patients

.02%

247

Completed Appointments

13%

78

Referrals

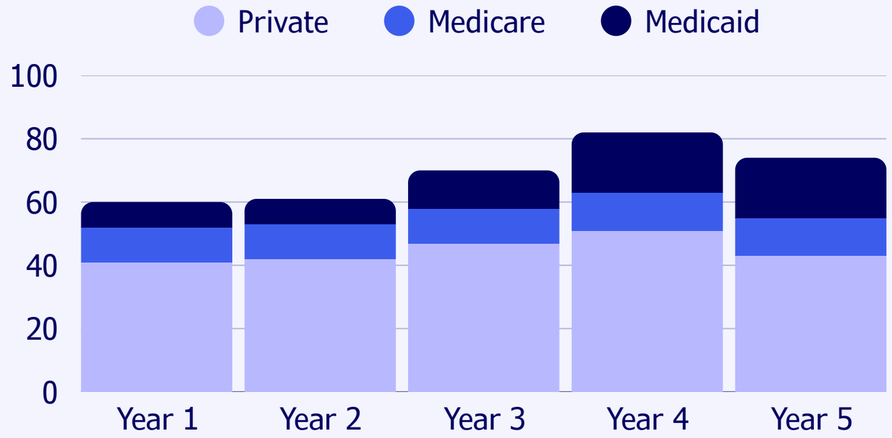
14%

50

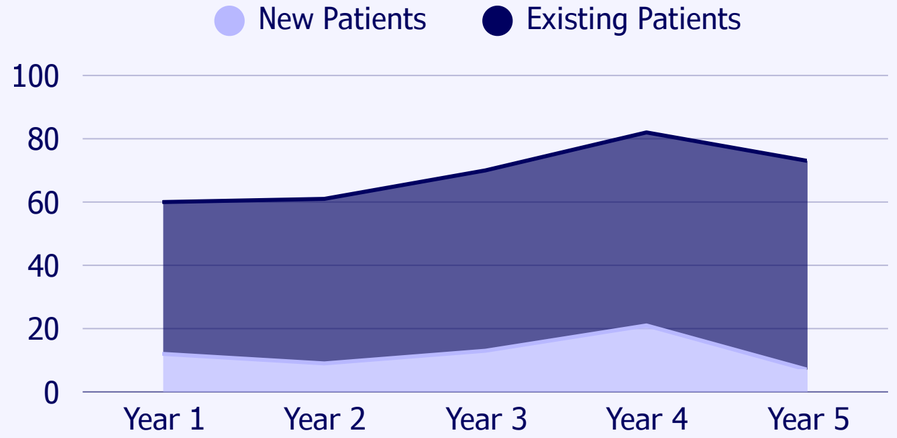
Online Consults

56%

32



Total Patient Population by Payor

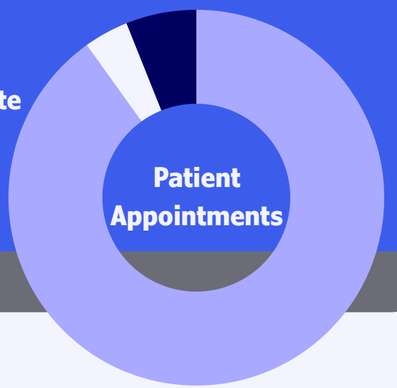


Total Patient Population Over the Most Recent 5 Years

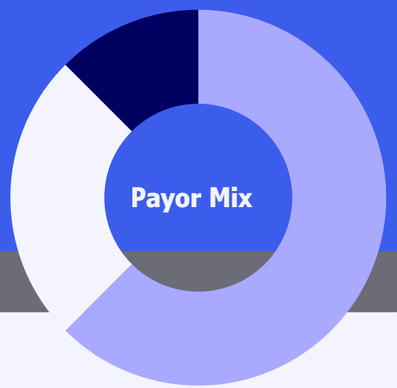
56% Total

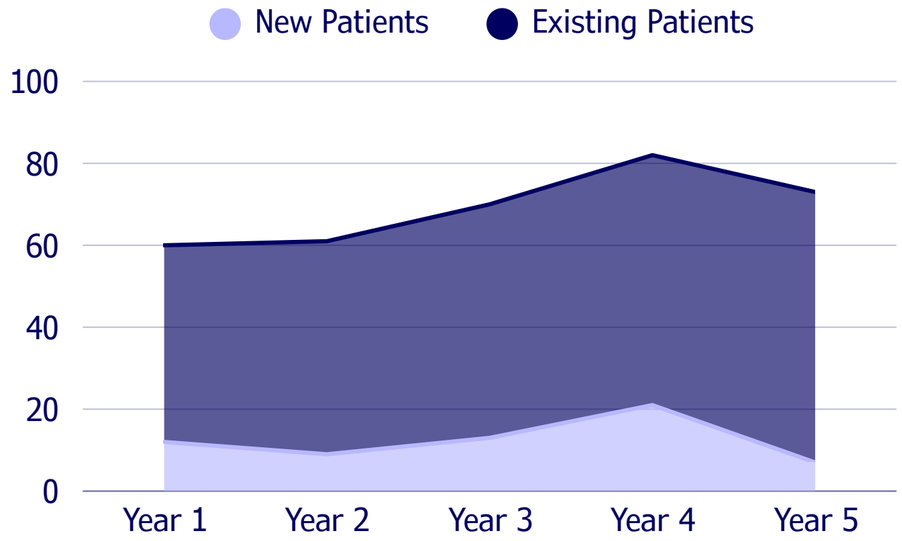
56% Total

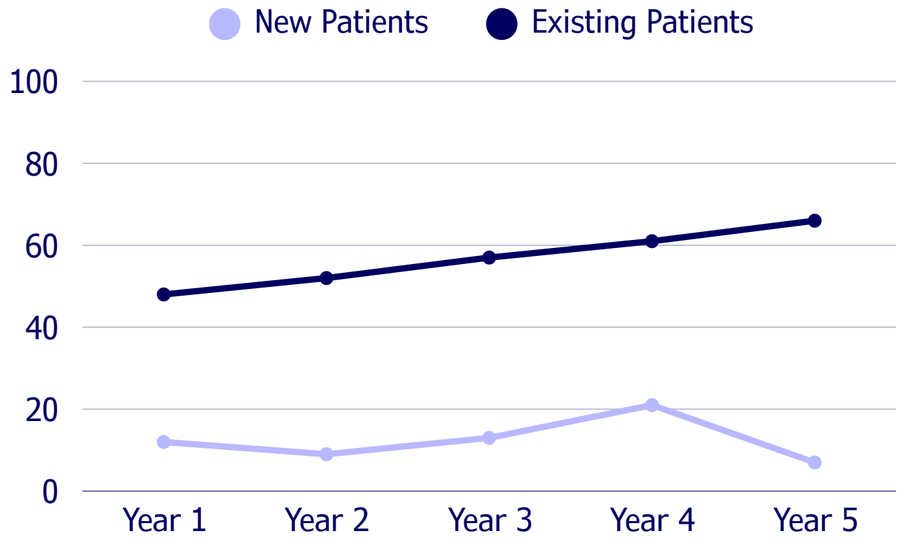
90.1%
Completion Rate

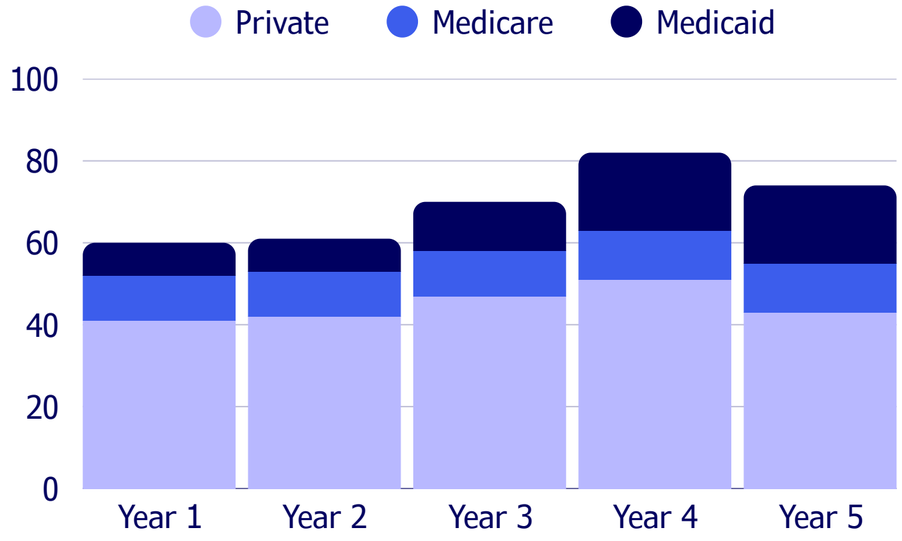


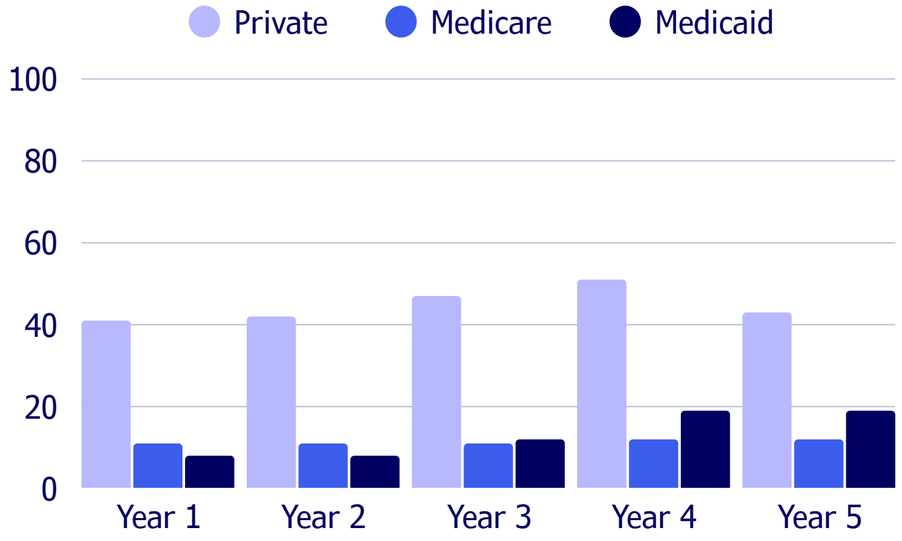
56% Total

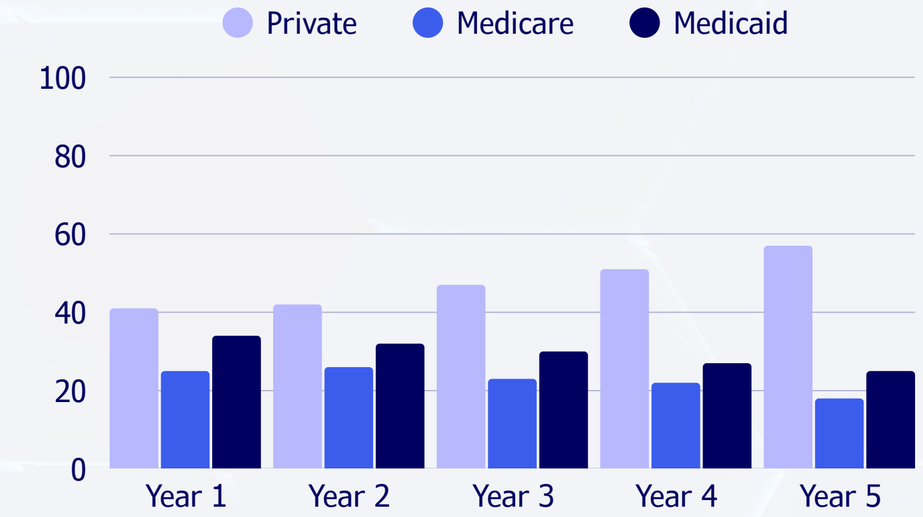


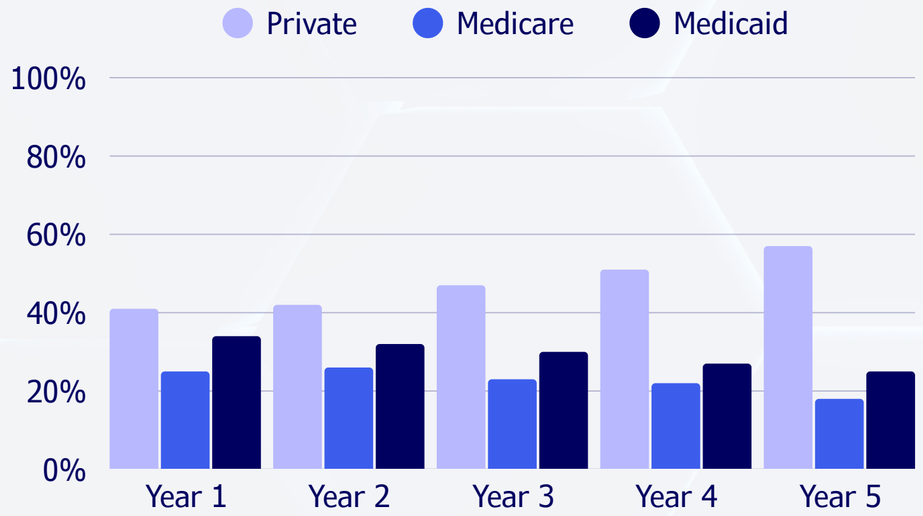

New Patients
Existing Patients
100
80
60
40
20
0
Year 1
Year 2
Year 3
Year 4
Year 5

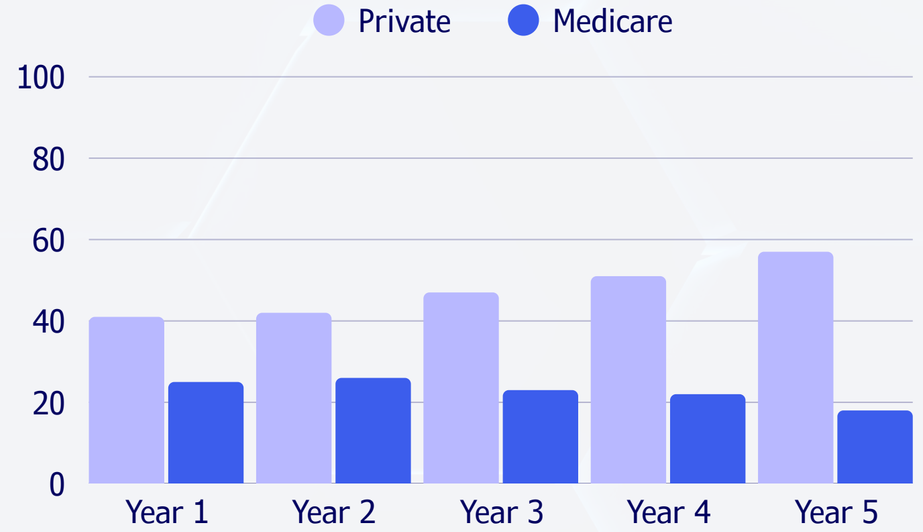

New Patients
Existing Patients
100
80
60
40
20
0
Year 1
Year 2
Year 3
Year 4
Year 5

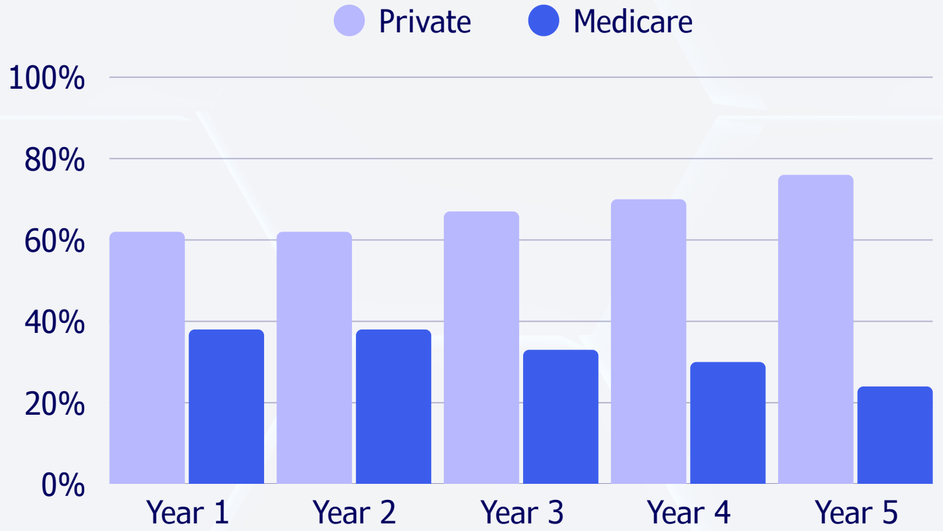

Private
Medicare
Medicaid
100
80
60
40
20
0
Year 1
Year 2
Year 3
Year 4
Year 5

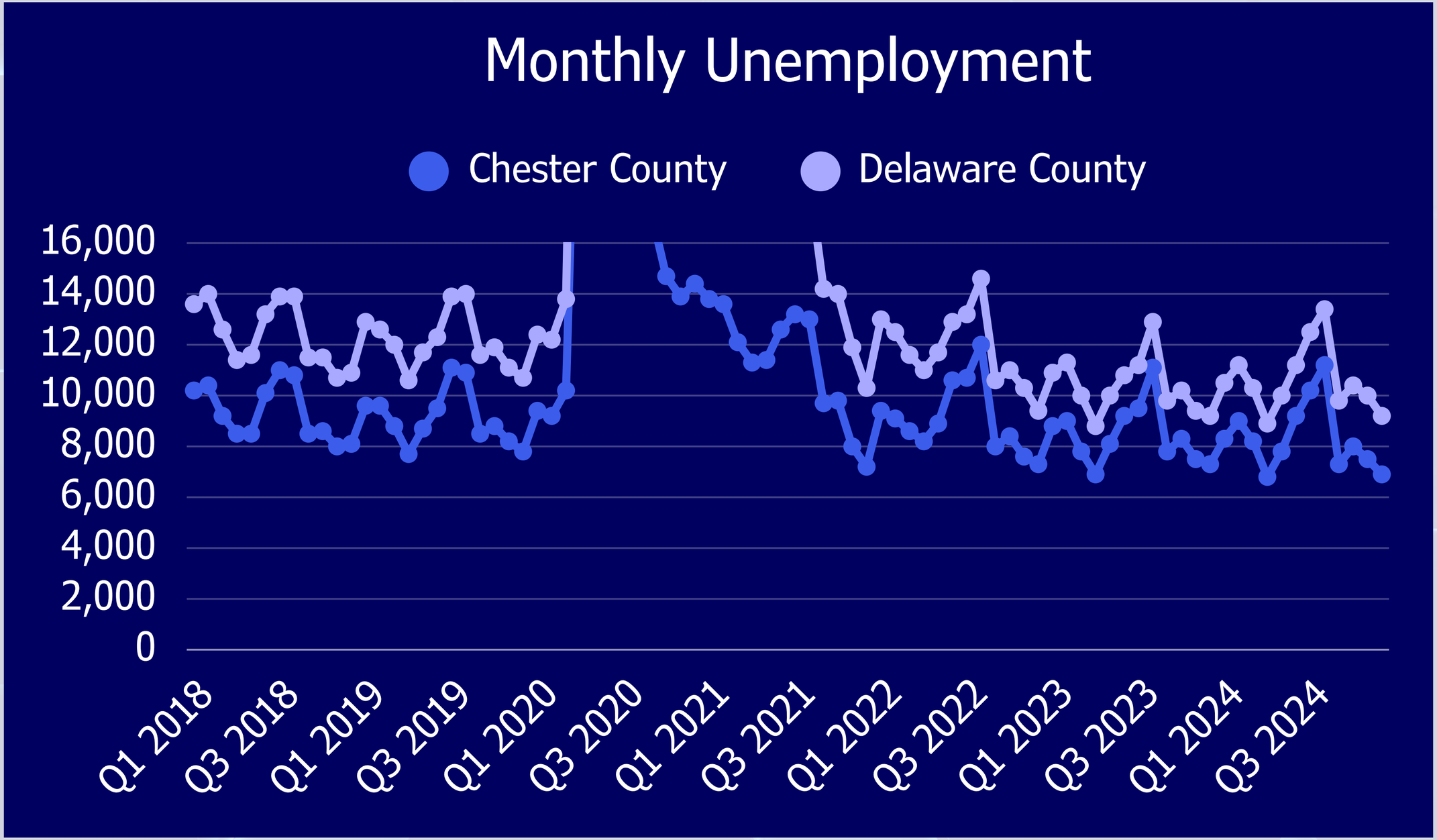

Private
Medicare
Medicaid
100
80
60
40
20
0
Year 1
Year 2
Year 3
Year 4
Year 5

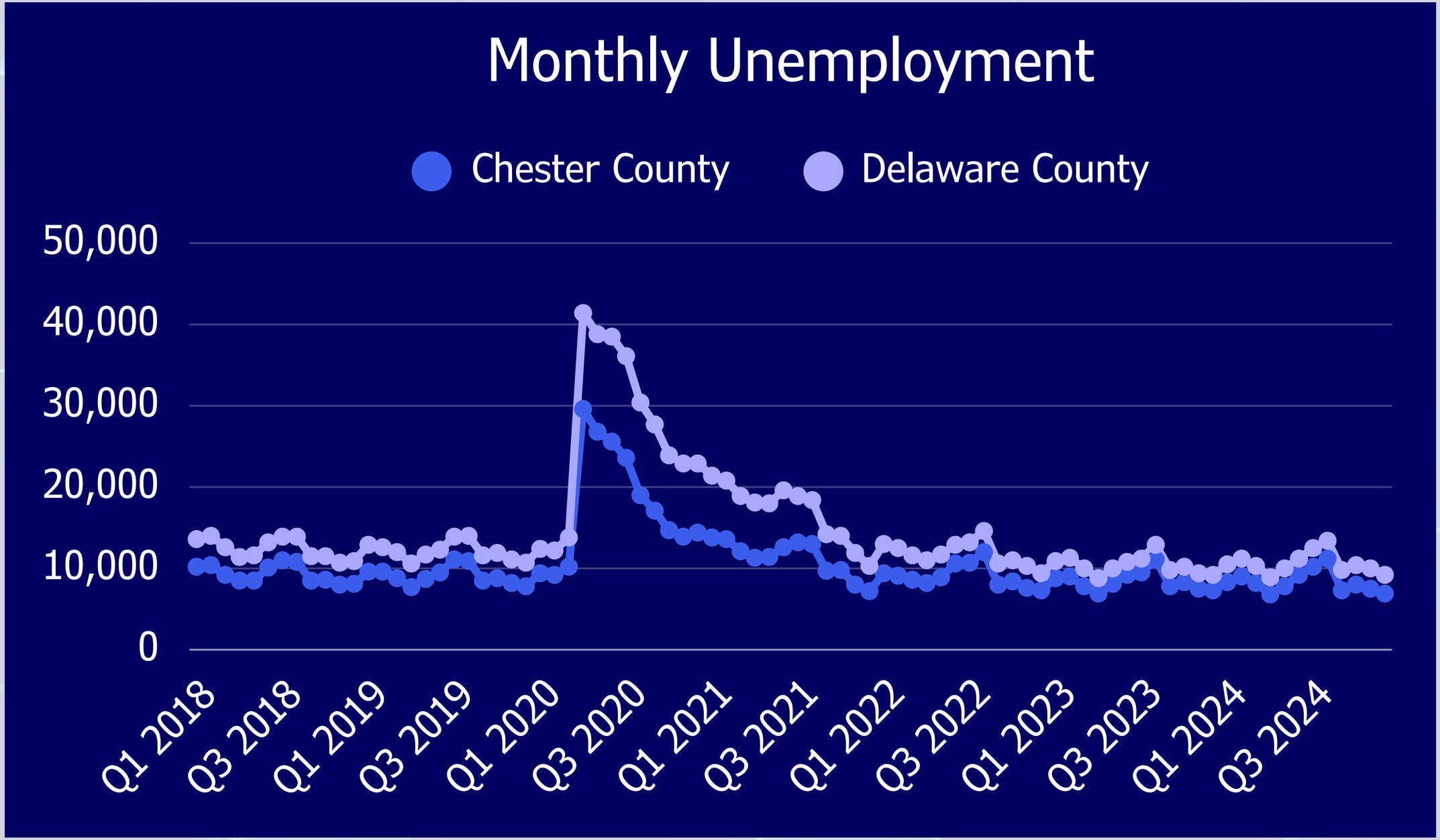

Private
Medicare
Medicaid
100
80
60
40
20
0
Year 1
Year 2
Year 3
Year 4
Year 5


Private
Medicare
Medicaid
100%
80%
60%
40%
20%
0%
Year 1
Year 2
Year 3
Year 4
Year 5

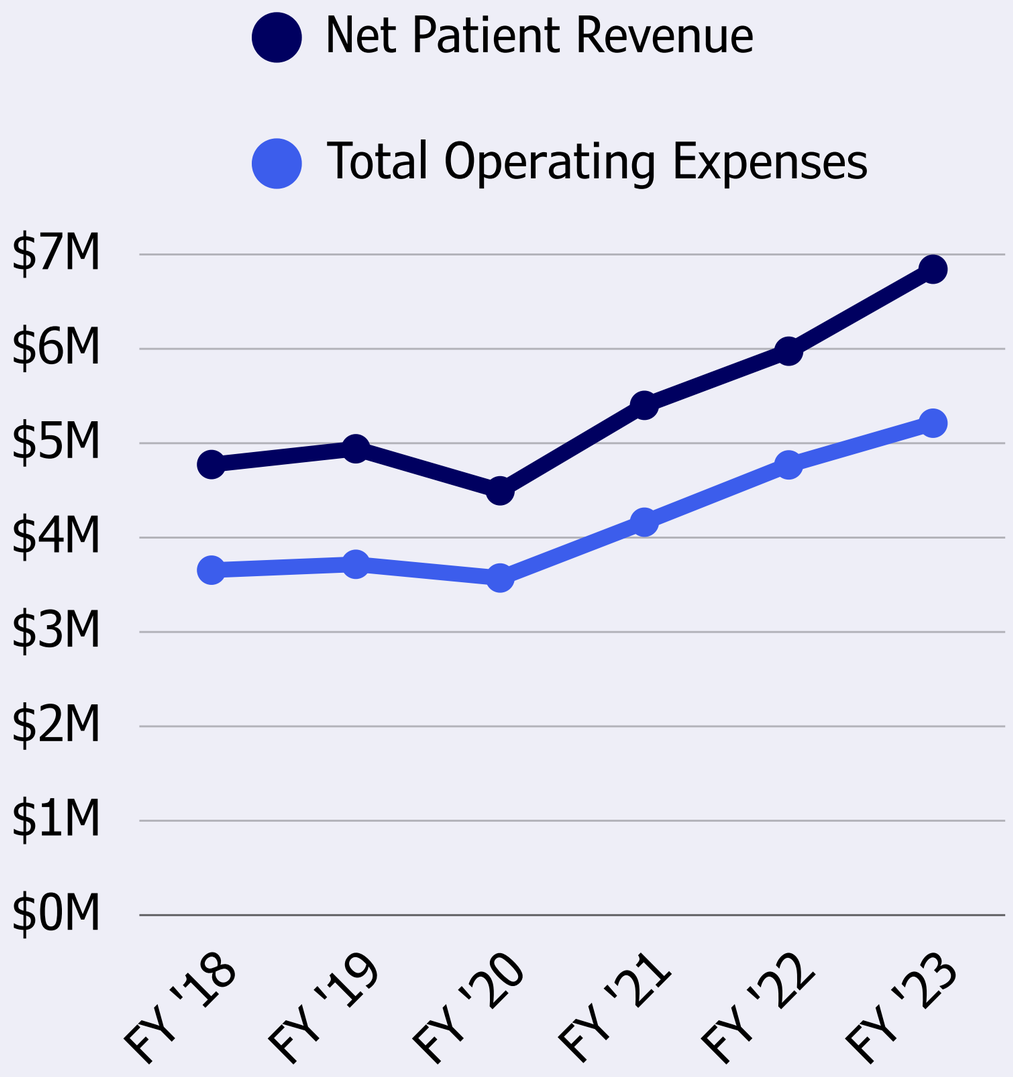

Private
Medicare
100
80
60
40
20
0
Year 1
Year 2
Year 3
Year 4
Year 5


Private
Medicare
100%
80%
60%
40%
20%
0%
Year 1
Year 2
Year 3
Year 4
Year 5


Monthly Unemployment
Chester County
Delaware County
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
Q1 2018
Q3 2018
Q1 2019
Q3 2019
Q1 2020
Q3 2020
Q1 2021
Q3 2021
Q1 2022
Q3 2022
Q1 2023
Q3 2023
Q1 2024
Q3 2024


Monthly Unemployment
Chester County
Delaware County
50,000
40,000
30,000
20,000
10,000
0
Q1 2018
Q3 2018
Q1 2019
Q3 2019
Q1 2020
Q3 2020
Q1 2021
Q3 2021
Q1 2022
Q3 2022
Q1 2023
Q3 2023
Q1 2024
Q3 2024

Statewide Average



Statewide Average

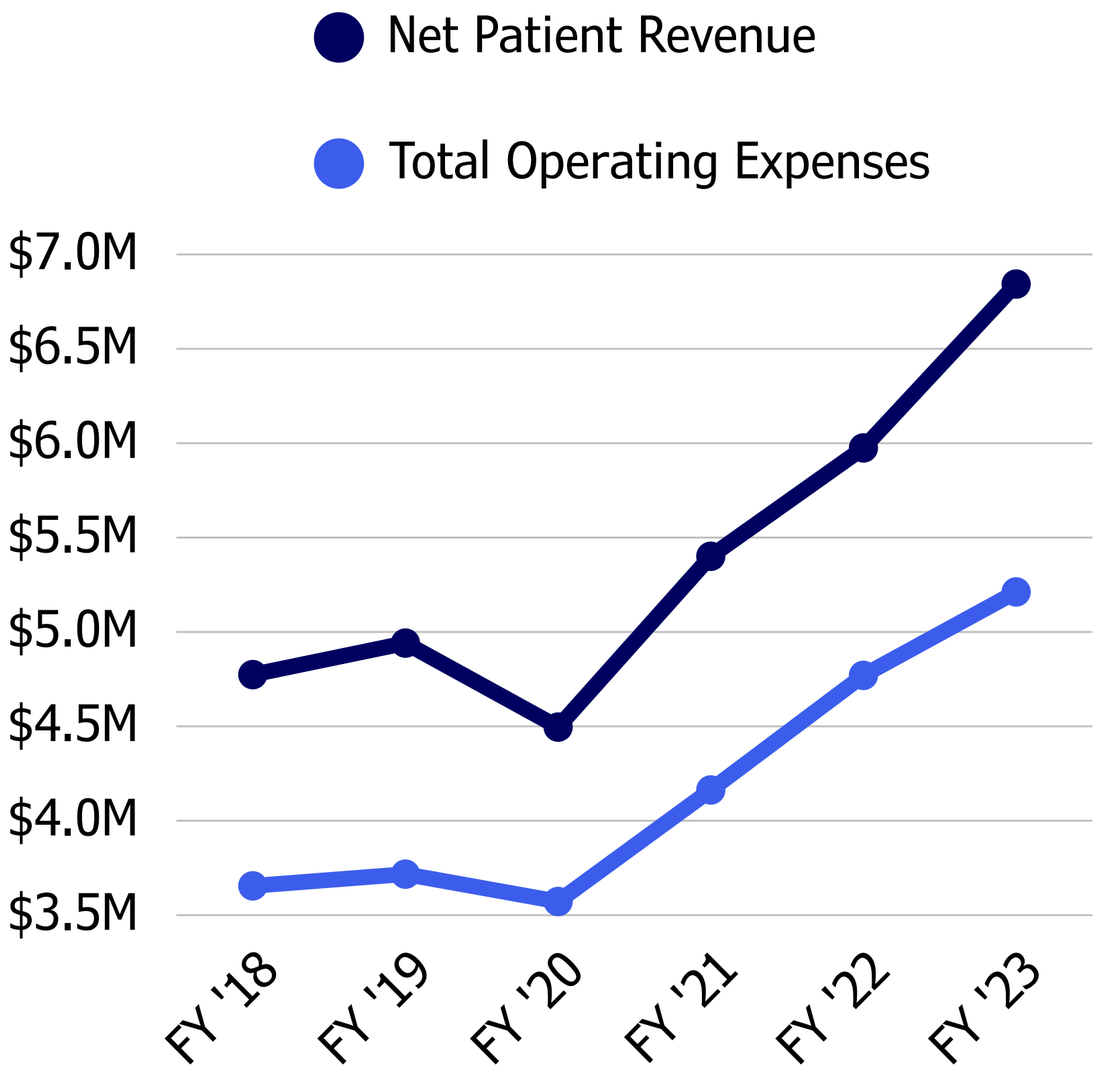
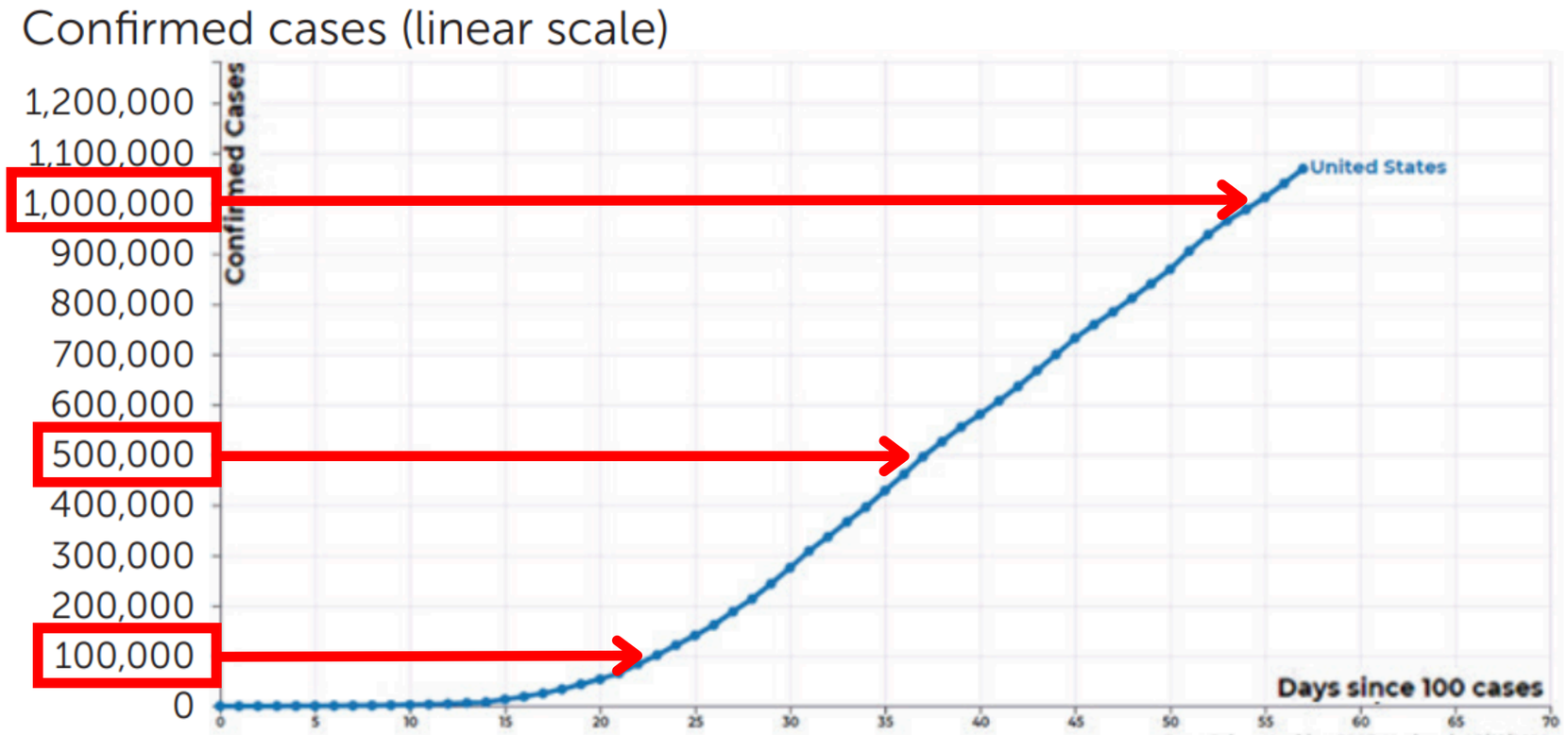
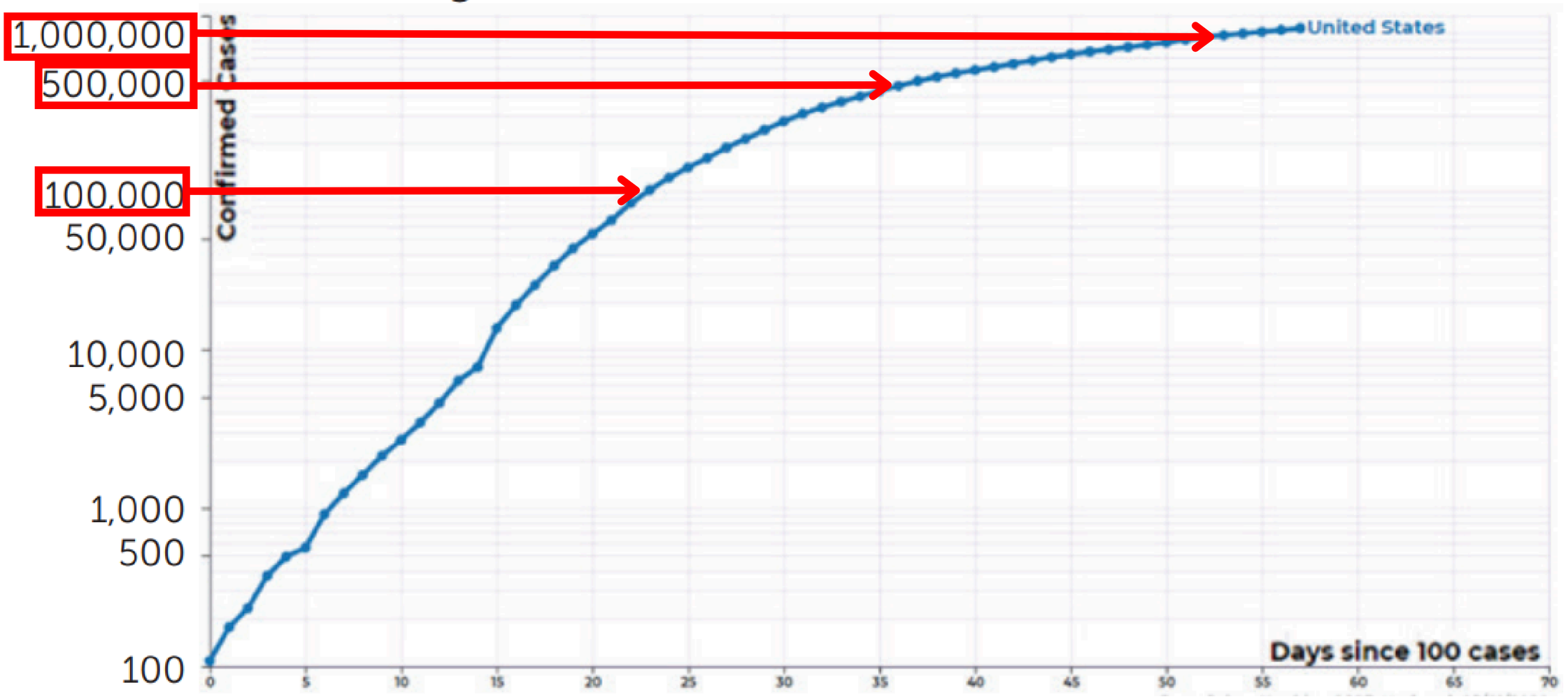


Figure 1. Linear & logarithmic graphs of COVID-19 cases in the United States as of May 1, 2020



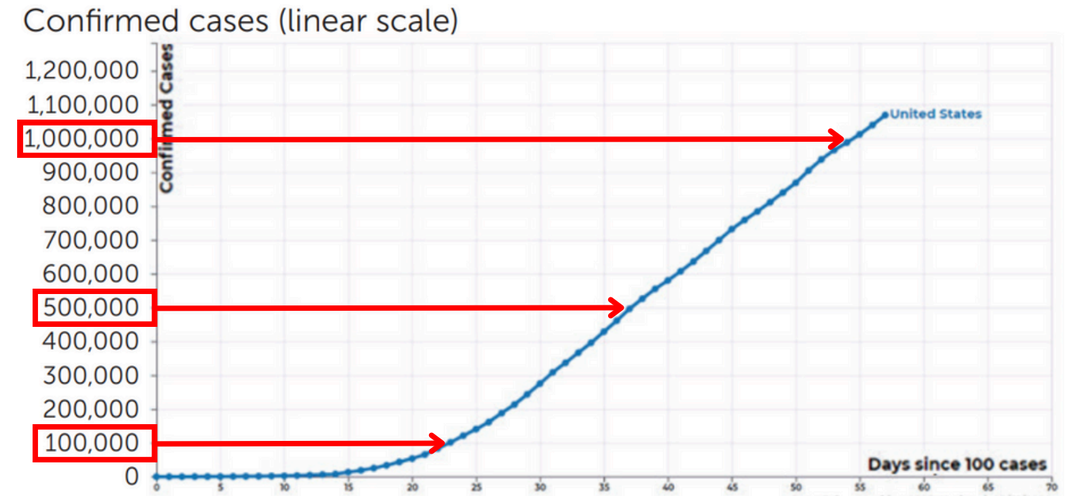
Citation:
Ryan, William H., & Evers, Ellen R. K. (2020). Graphs with logarithmic axes distort lay judgments. Behavioral Science & Policy. Retrieved from https://behavioralpolicy.org/ journal_issue/covid-19/

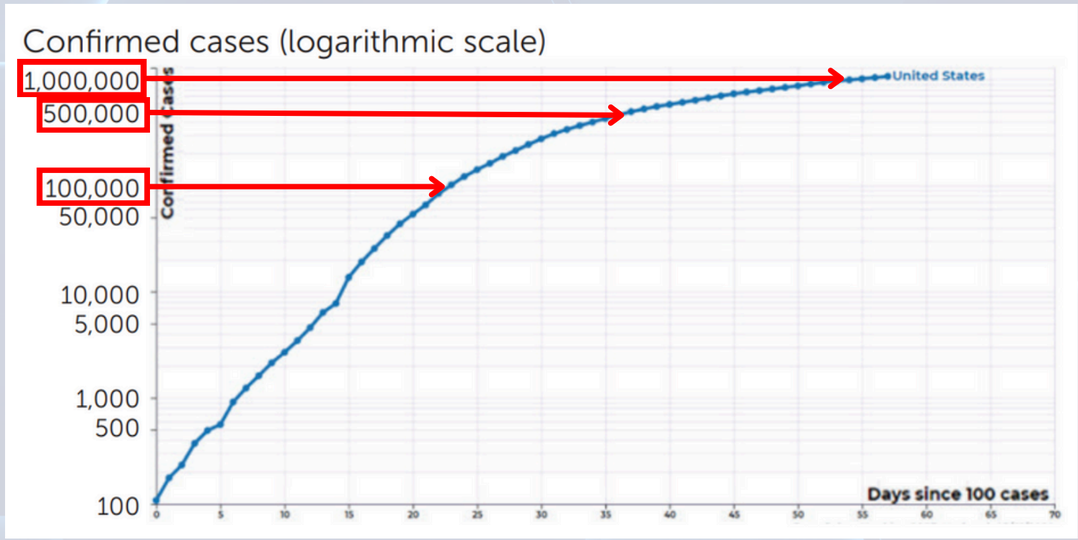
Confirmed cases (logarithmic scale)



Citation:
Ryan, William H., & Evers, Ellen R. K. (2020). Graphs with logarithmic axes distort lay judgments. Behavioral Science & Policy. Retrieved from https://behavioralpolicy.org/ journal_issue/covid-19/

Figure 1. Linear & logarithmic graphs of COVID-19 cases in the United States as of May 1, 2020





Citation:
Ryan, William H., & Evers, Ellen R. K. (2020). Graphs with logarithmic axes distort lay judgments. Behavioral Science & Policy. Retrieved from https://behavioralpolicy.org/ journal_issue/covid-19/

Health Care Organization Dashboard

Registered Patients

.02%

247

Completed Appointments

13%

78

Referrals

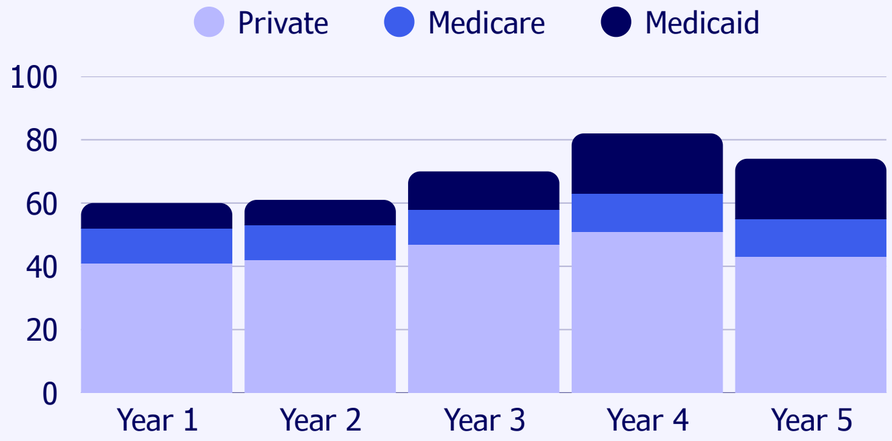
14%

50

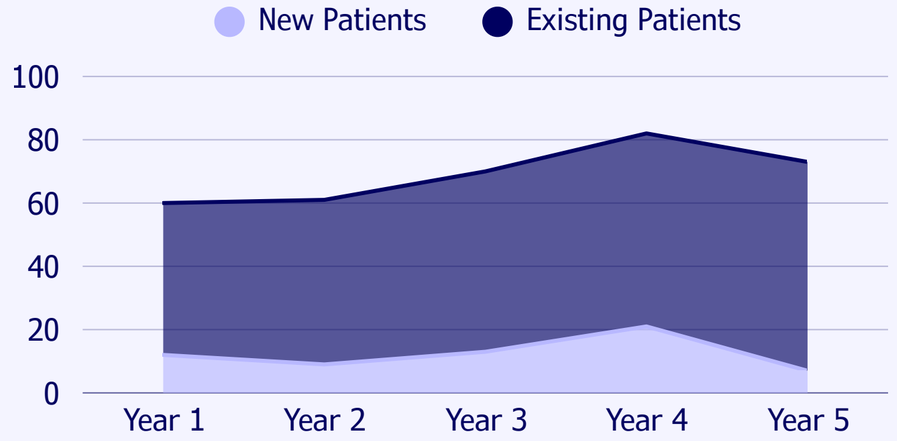
Online Consults

56%

32



Total Patient Population by Payor

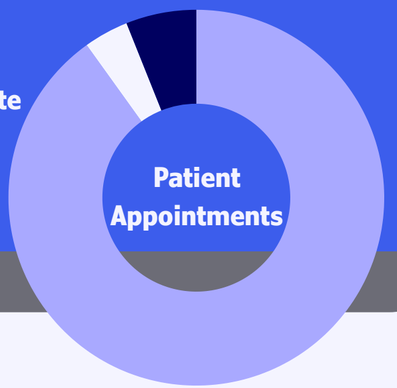


Total Patient Population Over the Most Recent 5 Years

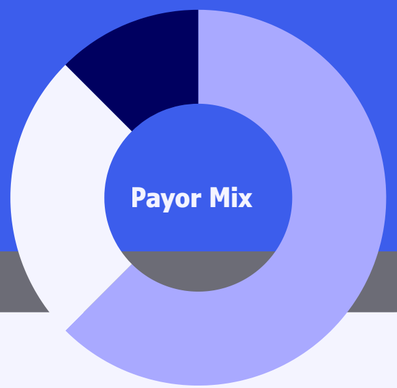
56% Total

56% Total

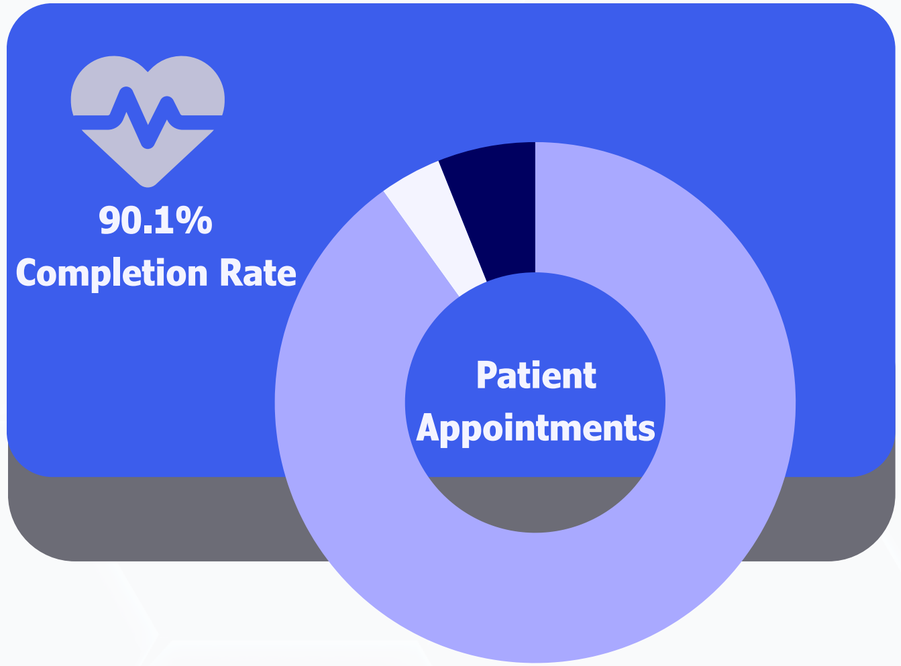
90.1%
Completion Rate



56% Total



Discussion Question



What are some questions you could ask of this data to better understand it?

PHC4
Transparent Health Care Reporting

Communicating Data

Factual

Origins, transformations and standards adhered to are accessible and easily verifiable

Transferable

Grounded in obtainable and suitable resources, transportable over populations, communities, and time

Ethical

Collected and communicated responsibly, respecting privacy, and avoiding misuse

Transparent

Conveying data insights through clear channels, with regard to community and professional standards

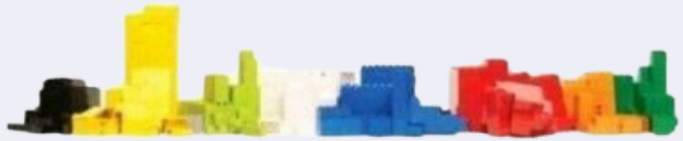
Communicating Data

PHC4
Transparent Health Care Reporting

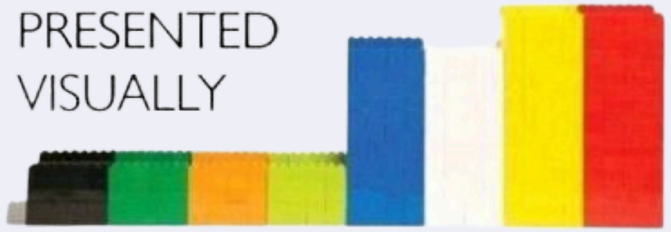
DATA

SORTED

ARRANGED



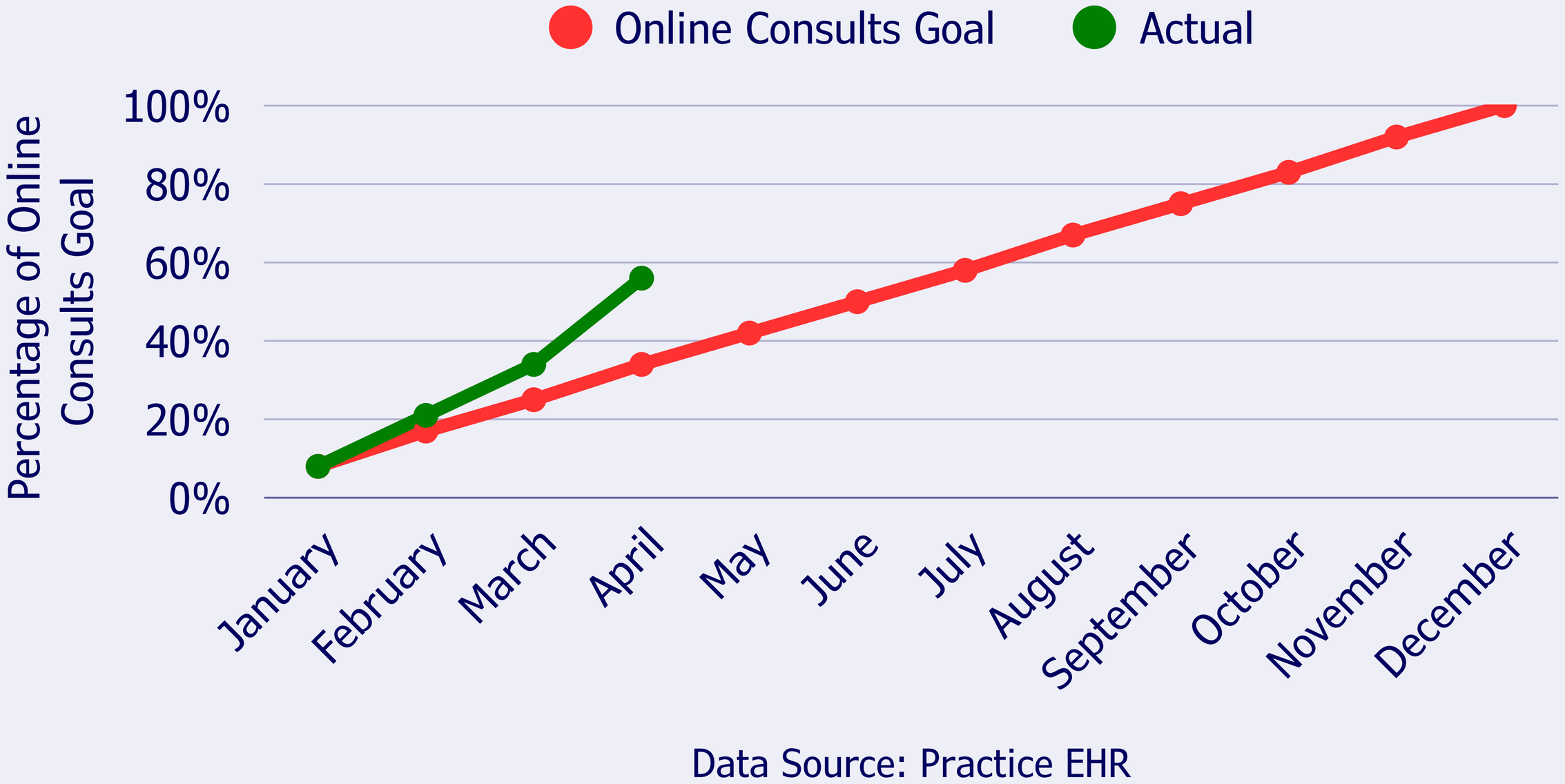
PRESENTED VISUALLY



EXPLAINED WITH A STORY

Credit:
Ryan Yockey

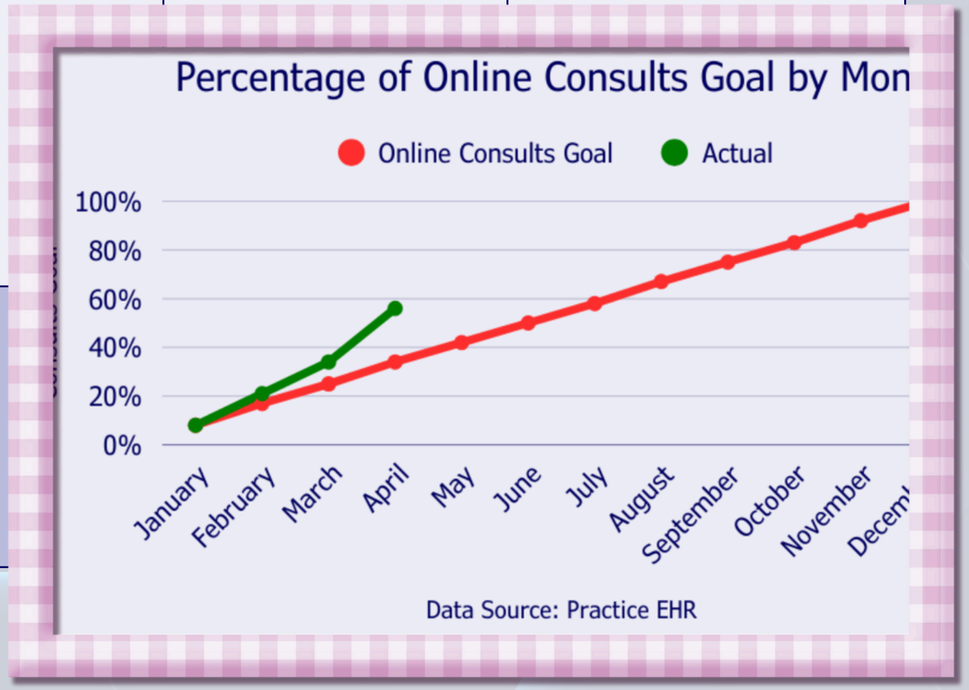
Percentage of Online Consults Goal by Month



Data Source: Practice EHR

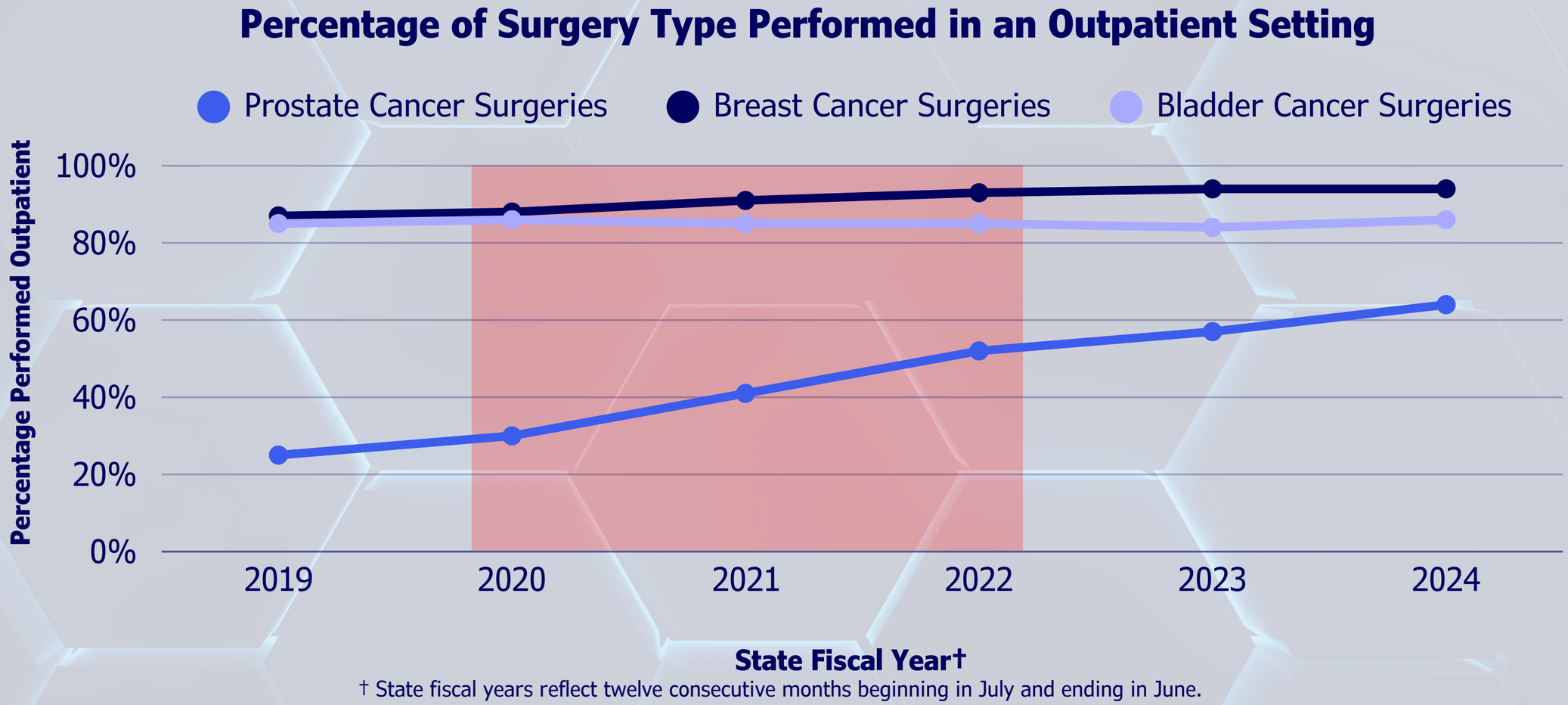
Details of Delivery

Data Source	Unit of Measurement	Target Audience	Desired Outcome	Setting	Discussion Type	Supporting Arguments
Online Consult Data from EHR	Percentage Increase in Online Consults Compared to Goal	Health Care Administrators	Increase Expected Revenue for Yearly Online Consults	Internal Director Meeting	Strategic Proposition	Time, ROI, Process Change



Communicating Visualizations and Data

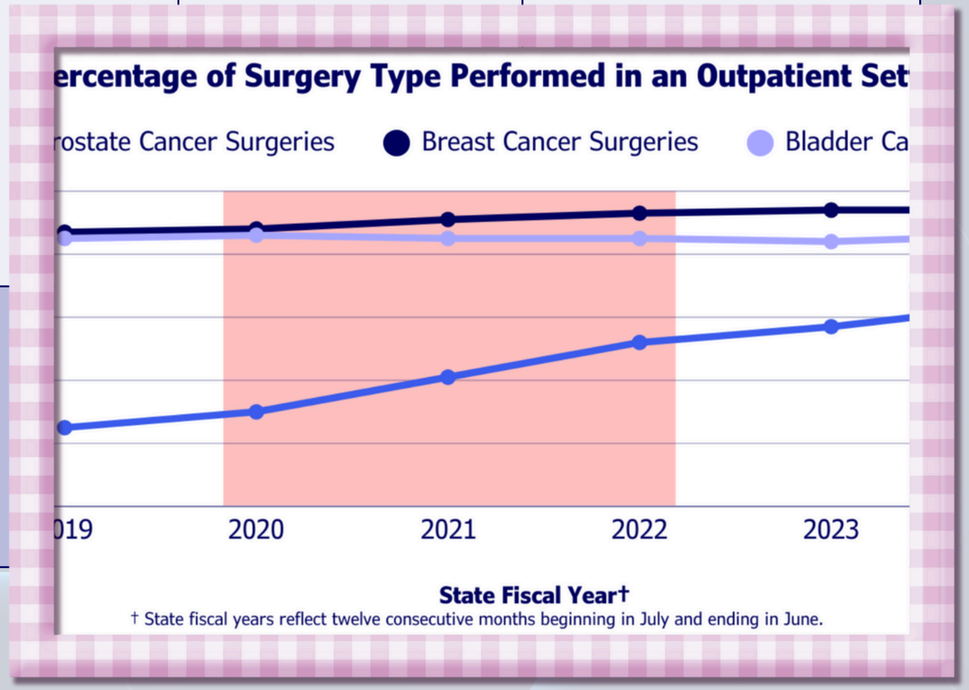
PHC4
Transparent Health Care Reporting

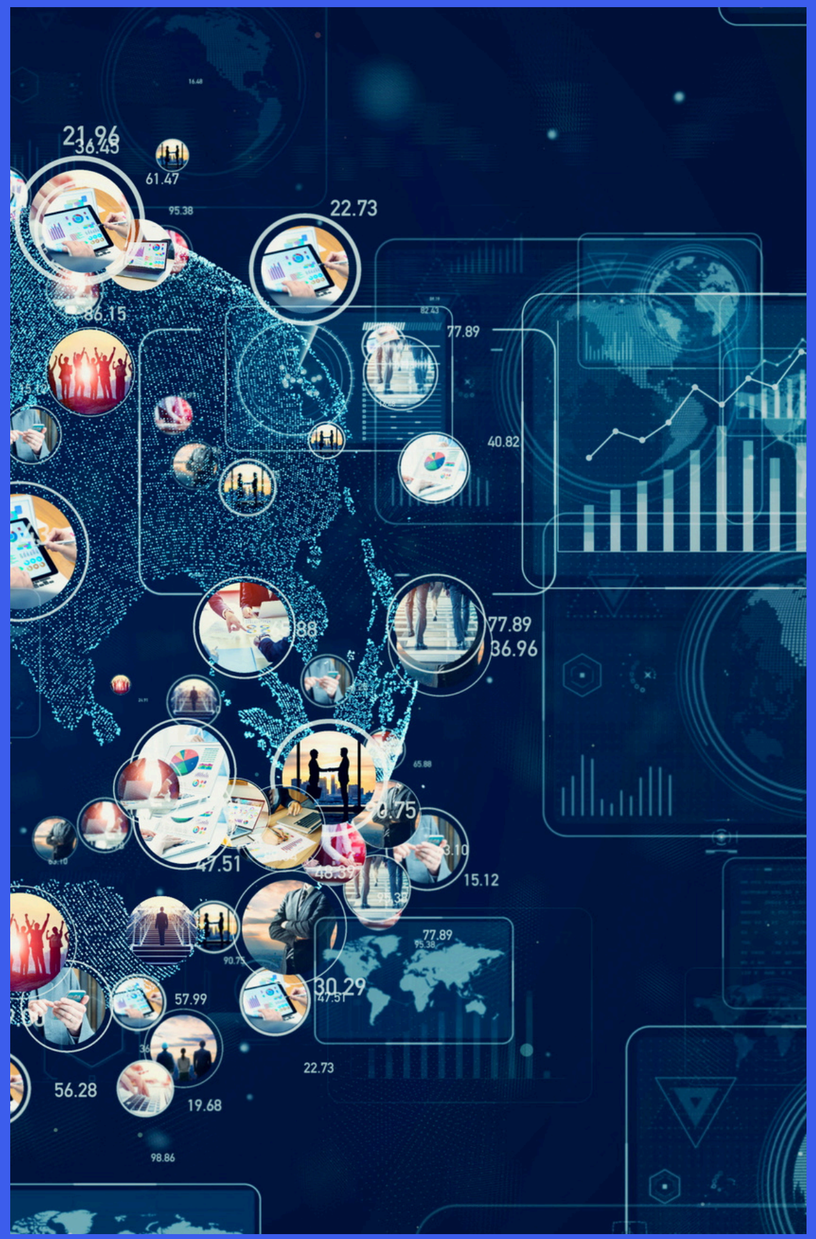


† State fiscal years reflect twelve consecutive months beginning in July and ending in June.

Details of Delivery

Data Source	Unit of Measurement	Target Audience	Desired Outcome	Setting	Discussion Type	Supporting Arguments
PHC4 Cancer Surgery Volume Report	Percentage of Procedures in the Outpatient during COVID	Health System Leaders	Cost Savings from Performing Surgery in a Lower Cost Setting	ACHE Meeting	Sharing Best Practices	Statewide Trends in 3 Types of Cancer Surgery





Thank You

Pennsylvania Health Care Cost Containment Council

PHC4
Transparent Health Care Reporting

Stay in touch