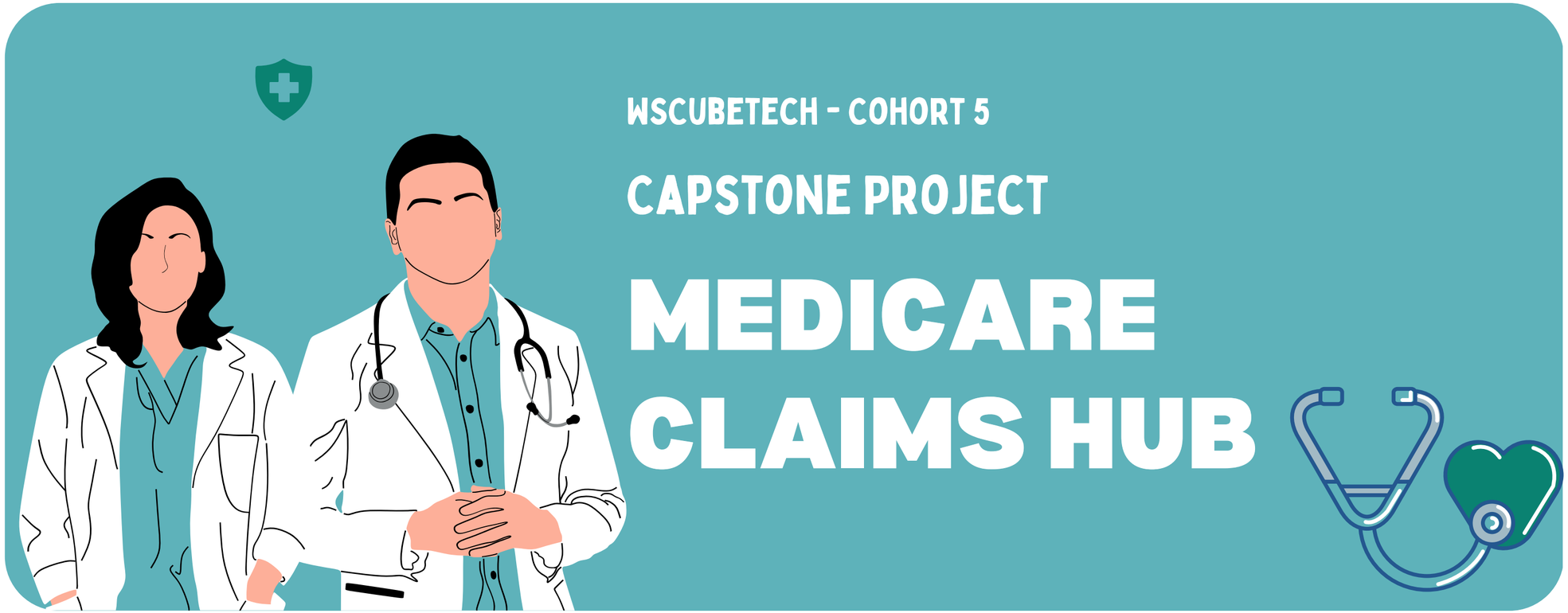


WSCUBETECH - COHORT 5

CAPSTONE PROJECT

MEDICARE CLAIMS HUB

Binay Kumar Naik

14-08-2025

DATA ANALYTICS & MACHINE LEARNING AT MEDICARE CLAIMS HUB

EXPLORATION OF INNOVATIONS IN MEDICINE

As a Data Analyst at MediCare, United States, your mission is to leverage claims data to uncover actionable insights, drive operational efficiencies, and build models to detect fraudulent activities. This capstone project simulates real-world analytics challenges in the healthcare claims management domain, where data-driven decisions can directly impact patient care quality and organizational profitability.

PHASE 1: ADVANCED SQL ANALYSIS – HEALTHCARE CLAIMS

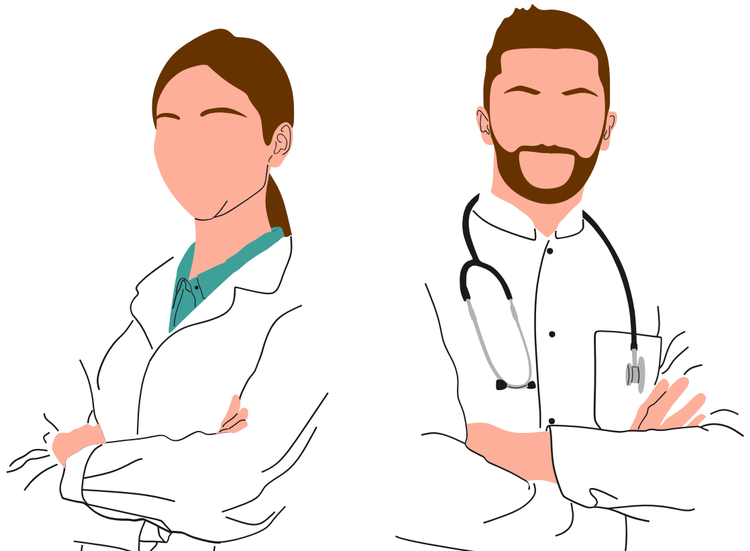
ADVANCED SQL ANALYSIS – HEALTHCARE CLAIMS

Use SQL to perform deep analysis of healthcare claims data, answering strategic questions for MediCare Claims Hub to better manage reimbursements, monitor provider performance, understand demographics, and identify patterns associated with chronic conditions.

Q1. Retrieve the total amount reimbursed for inpatient claims (InscClaimAmtReimbursed), grouped by provider.

QUERRY

```
SELECT provider,
SUM(InscClaimAmtReimbursed) AS total_amount
FROM inpatientdata
GROUP BY provider;
```



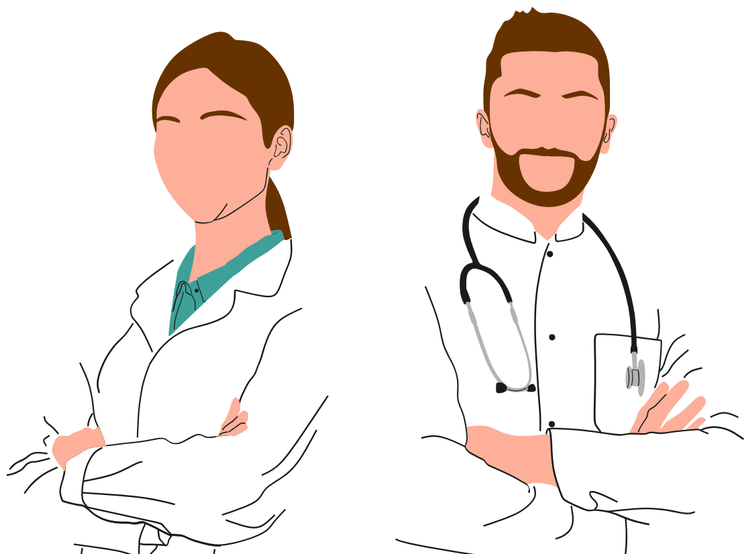
OUTPUT

provider	total_amount
PRV57070	195000
PRV54750	78000
PRV53758	1173900
PRV55825	789600
PRV52338	1564900
PRV55544	111000
PRV53275	1456000
PRV54989	423000

Q2. Could you identify the top 5 providers with the highest number of outpatient claims?

QUERRY

```
SELECT provider,
COUNT(ClaimID) as claim_count
FROM outpatientdata
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5;
```



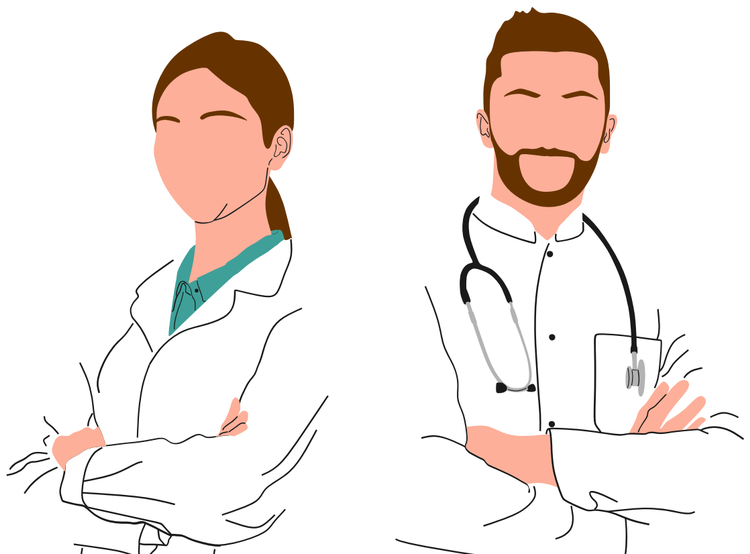
OUTPUT

	provider	claim_count
▶	PRV56573	3065
	PRV52080	3036
	PRV55485	2823
	PRV53105	2763
	PRV51939	1804

Q3. Find the total number of beneficiaries with claims indicating chronic conditions such as diabetes (ChronicCond_Diabetes = 1)

QUERRY

```
SELECT COUNT(BeneID) as
Total_Claims_Diabetes
FROM beneficiarydata
WHERE ChronicCond_Diabetes = 1;
```



OUTPUT

Total_Claims_Diabetes
41786

Q4. Calculate the average inpatient claim amount reimbursed by gender.

QUERRY

```
SELECT Gender,
AVG(ip.InscClaimAmtReimbursed) as
Avg_Claim_Amount
FROM inpatientdata as ip
JOIN beneficiarydata as bf
WHERE ip.BeneID = bf.BeneID
GROUP BY Gender;
```

OUTPUT

Gender	Avg_Claim_Amount
2	10112.5308
1	10095.7421

Q5. Retrieve all claims (inpatient & outpatient) for a given BeneID, to enable individual beneficiary case history reviews.

QUERRY

```
SELECT BeneID, ClaimID, ClaimStartDt, ClaimEndDt, Provider, InscClaimAmtReimbursed
FROM inpatientdata
WHERE BeneID = 'BENE11725'
UNION
SELECT BeneID, ClaimID, ClaimStartDt, ClaimEndDt, Provider, InscClaimAmtReimbursed
FROM outpatientdata
WHERE BeneID = 'BENE11725';
```

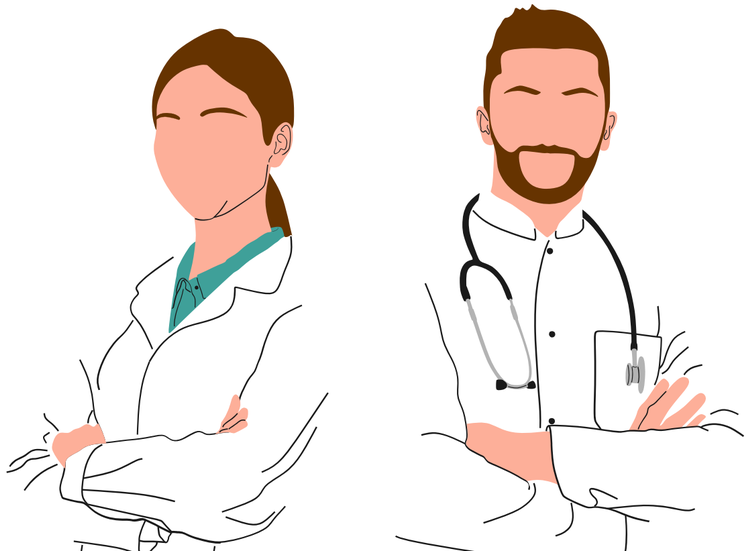
OUTPUT

BeneID	ClaimID	ClaimStartDt	ClaimEndDt	Provider	InscClaimAmtReimbursed
BENE11725	CLM37792	2009-02-10	2009-02-11	PRV55204	3000
BENE11725	CLM170452	2009-02-01	2009-02-01	PRV55204	80

Q6. Identify providers with claims where the admission date is in 2009 and the reimbursed amount exceeds $10,000.

QUERRY

```
SELECT Provider, AdmissionDt, InscClaimAmtReimbursed
FROM inpatientdata
WHERE YEAR(AdmissionDt) = 2009
AND InscClaimAmtReimbursed > 10000;
```



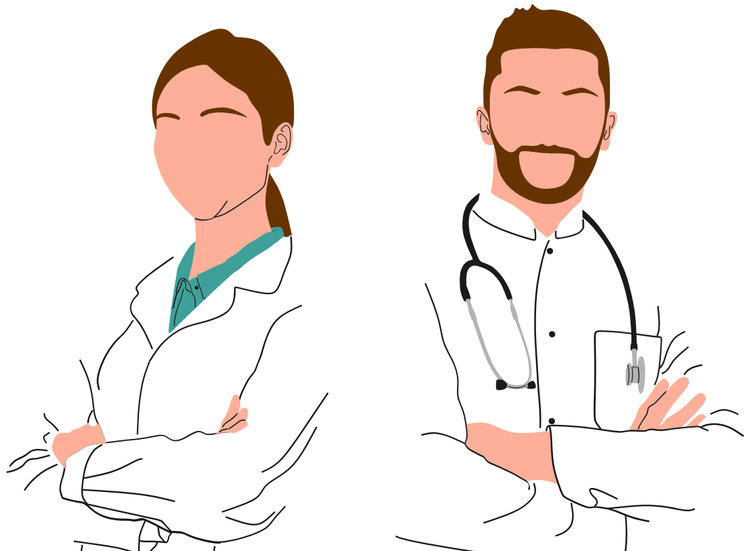
OUTPUT

Provider	AdmissionDt	InscClaimAmtReimbursed
PRV55825	2009-06-23	16000
PRV52338	2009-01-20	19000
PRV57214	2009-06-14	15000
PRV52117	2009-08-24	19000
PRV54853	2009-03-02	11000
PRV53275	2009-07-11	13000
PRV52091	2009-05-30	16000
PRV54731	2009-04-05	25000
PRV53866	2009-05-03	11000
PRV51085	2009-03-15	14000

Q7. Combine beneficiary demographics with inpatient claims to calculate the average deductible amount (IPAnnualDeductibleAmt) for beneficiaries aged 65 and above.

QUERRY

```
SELECT AVG(DeductibleAmtPaid) AS Avg_deductible
FROM inpatientdata AS ip
JOIN beneficiarydata AS bf
ON ip.BeneID = bf.BeneID
WHERE TIMESTAMPDIFF(YEAR, bf.DOB, CURDATE()) >= 65;
```



OUTPUT

Avg_deductible
1068

Q8. List all claims involving more than one physician (i.e., where AttendingPhysician, OperatingPhysician, or OtherPhysician are non-null simultaneously).

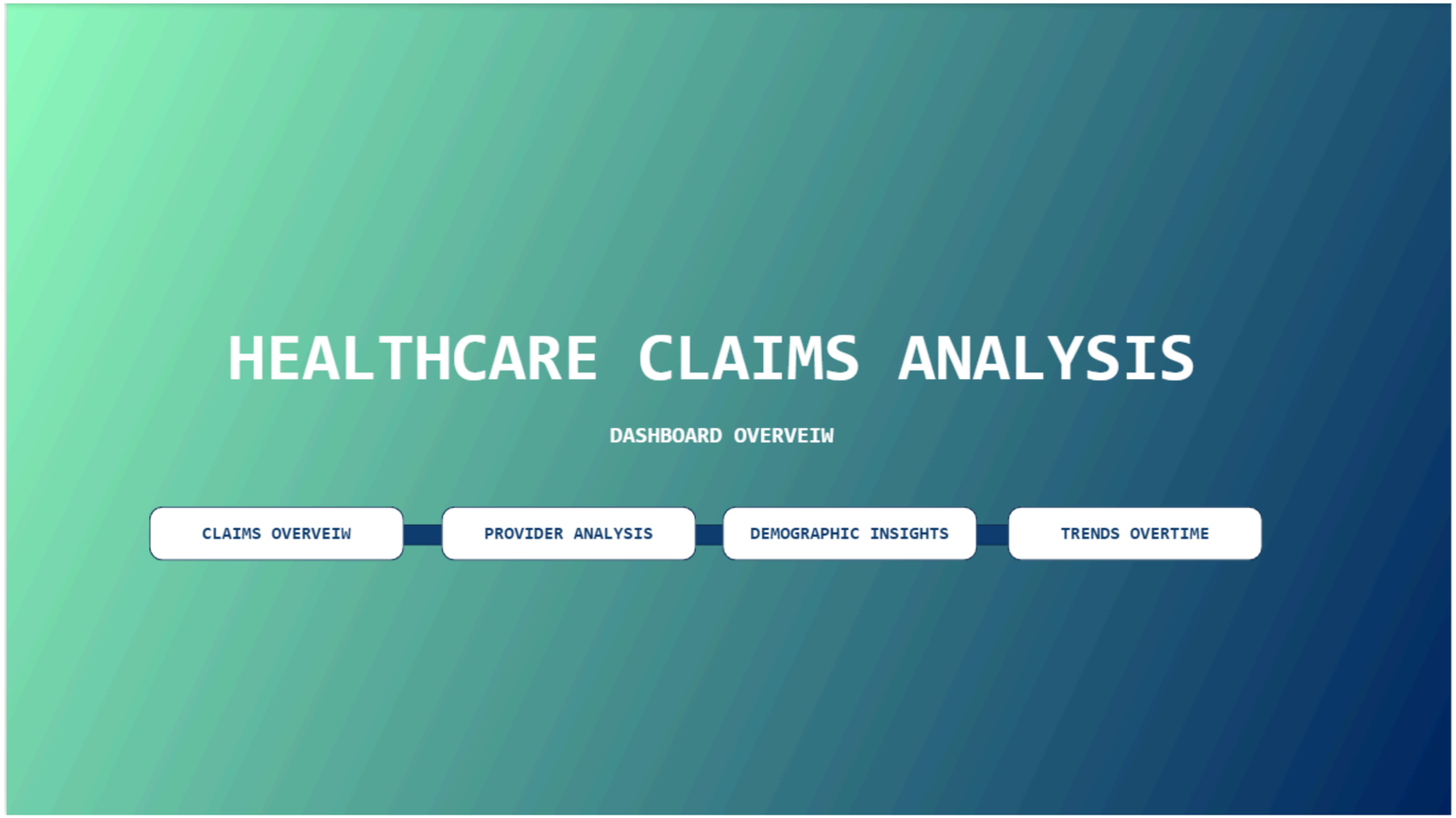
QUERRY

```
SELECT BeneID, ClaimID, ClaimStartDt, ClaimEndDt, Provider, InscClaimAmtReimbursed, AttendingPhysician,
OperatingPhysician, OtherPhysician
FROM inpatientdata
WHERE (
(AttendingPhysician IS NOT NULL AND AttendingPhysician <> '')
+(OperatingPhysician IS NOT NULL AND OperatingPhysician <> '')
+(OtherPhysician IS NOT NULL AND OtherPhysician <> '')) >= 2

UNION

SELECT BeneID, ClaimID, ClaimStartDt, ClaimEndDt, Provider, InscClaimAmtReimbursed, AttendingPhysician,
OperatingPhysician, OtherPhysician
FROM outpatientdata
WHERE (
(AttendingPhysician IS NOT NULL AND AttendingPhysician <> '')
+(OperatingPhysician IS NOT NULL AND OperatingPhysician <> '')
+(OtherPhysician IS NOT NULL AND OtherPhysician <> '')) >= 2;
```

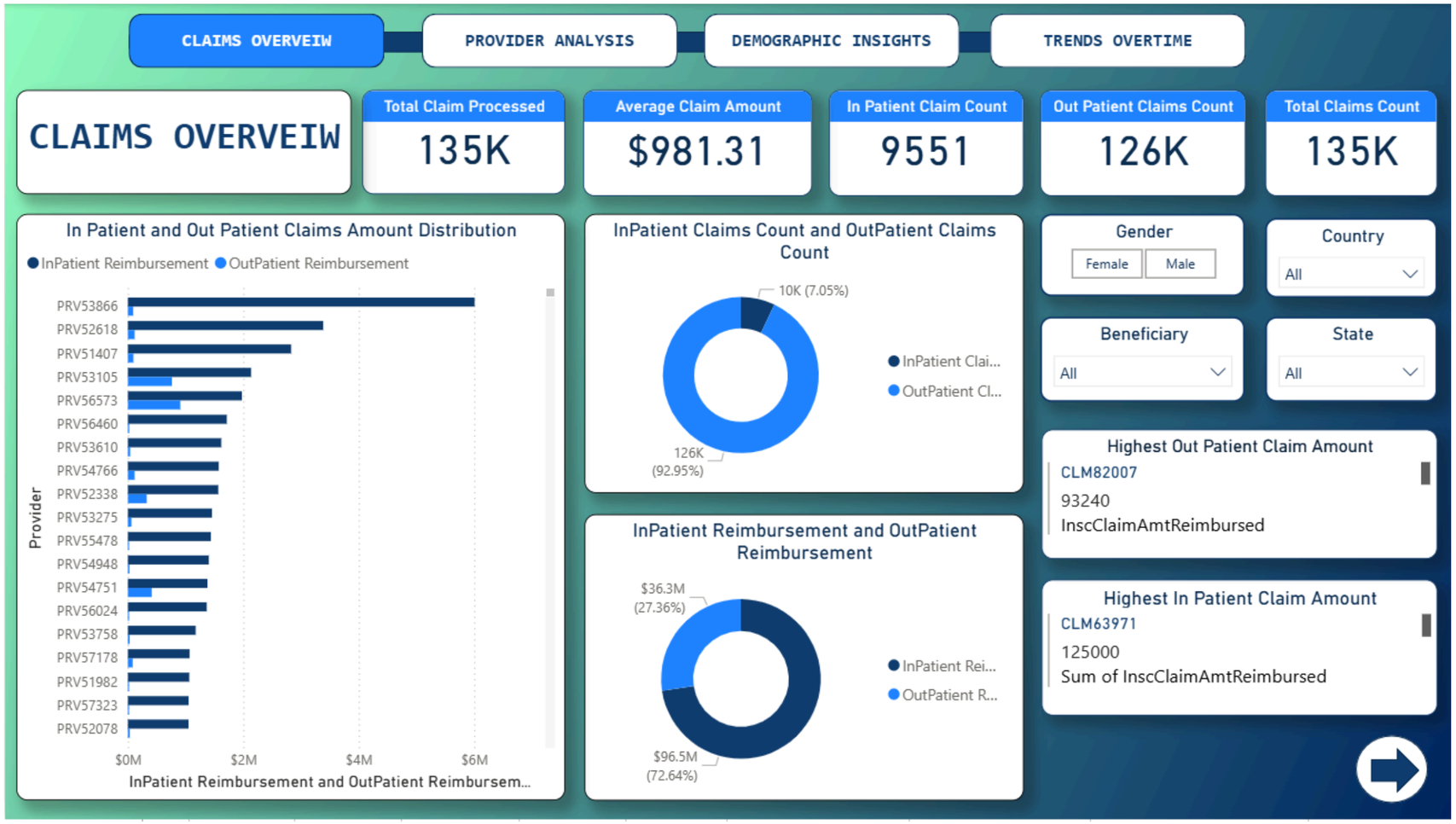
PHASE 2: INTERACTIVE DASHBOARDS – POWER BI OR TABLEAU

- Create a comprehensive, interactive dashboard for MediCare Claims Hub leadership to visualize claims patterns, demographic impacts, provider performance, and time-based trends, empowering data-driven decision-making.

HOME PAGE


HEALTHCARE CLAIMS ANALYSIS
DASHBOARD OVERVEIW
CLAIMS OVERVEIW
PROVIDER ANALYSIS
DEMOGRAPHIC INSIGHTS
TRENDS OVERTIME

CLAIMS OVERVEIW



CLAIMS OVERVEIW

Over 135,000 claims were processed, averaging $981 each.

The split tells two very different stories:

- Outpatient claims dominate – 93% of total, worth $96.5M.
- Inpatient claims, though just 7%, bring $36.3M, highlighting their higher value per case.

Notable outliers add more color:

- Largest outpatient claim: $93,240 (CLM82007).
- Largest inpatient claim: $125,000 (CLM63971).

Across providers, some like PRV53866 process significantly higher amounts, yet overall the system runs steady, near $1,000 per claim.

PROVIDER ANALYSIS

PROVIDER ANALYSIS

Over 1,353 providers were active, managing $132.9M in reimbursements.

The financial weight skews heavily:

- Inpatient reimbursements = $96.5M.
- Outpatient reimbursements = $36.3M.

A few providers clearly lead the pack:

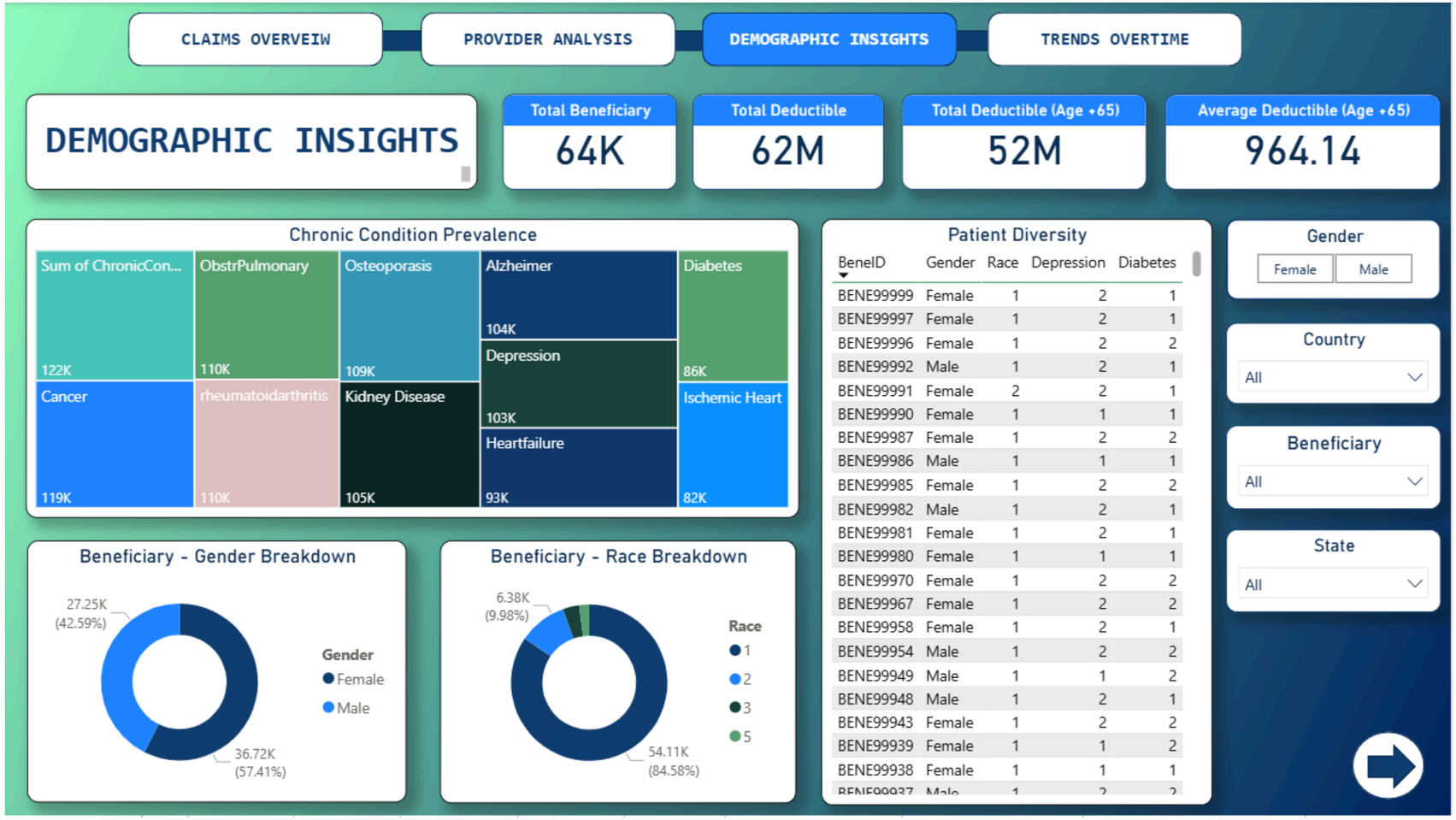
- PRV53866 tops the list with nearly $6M.
- PRV52618, PRV51407, PRV53105, and PRV56573 each handle ~ $3M.

When it comes to claim volumes:

- PRV56573 and PRV52080 lead for outpatient.
- PRV53866 records the highest inpatient activity.

A small core group of providers shoulders a large share of reimbursements—key influencers in the system.

DEMOGRAPHIC ANALYSIS



DEMOGRAPHIC ANALYSIS

More than 64,000 beneficiaries are covered, with $62M in total deductibles.

Age drives the difference:

- Seniors 65+ contribute $52M in deductibles, averaging $964 per person.

The population snapshot is telling:

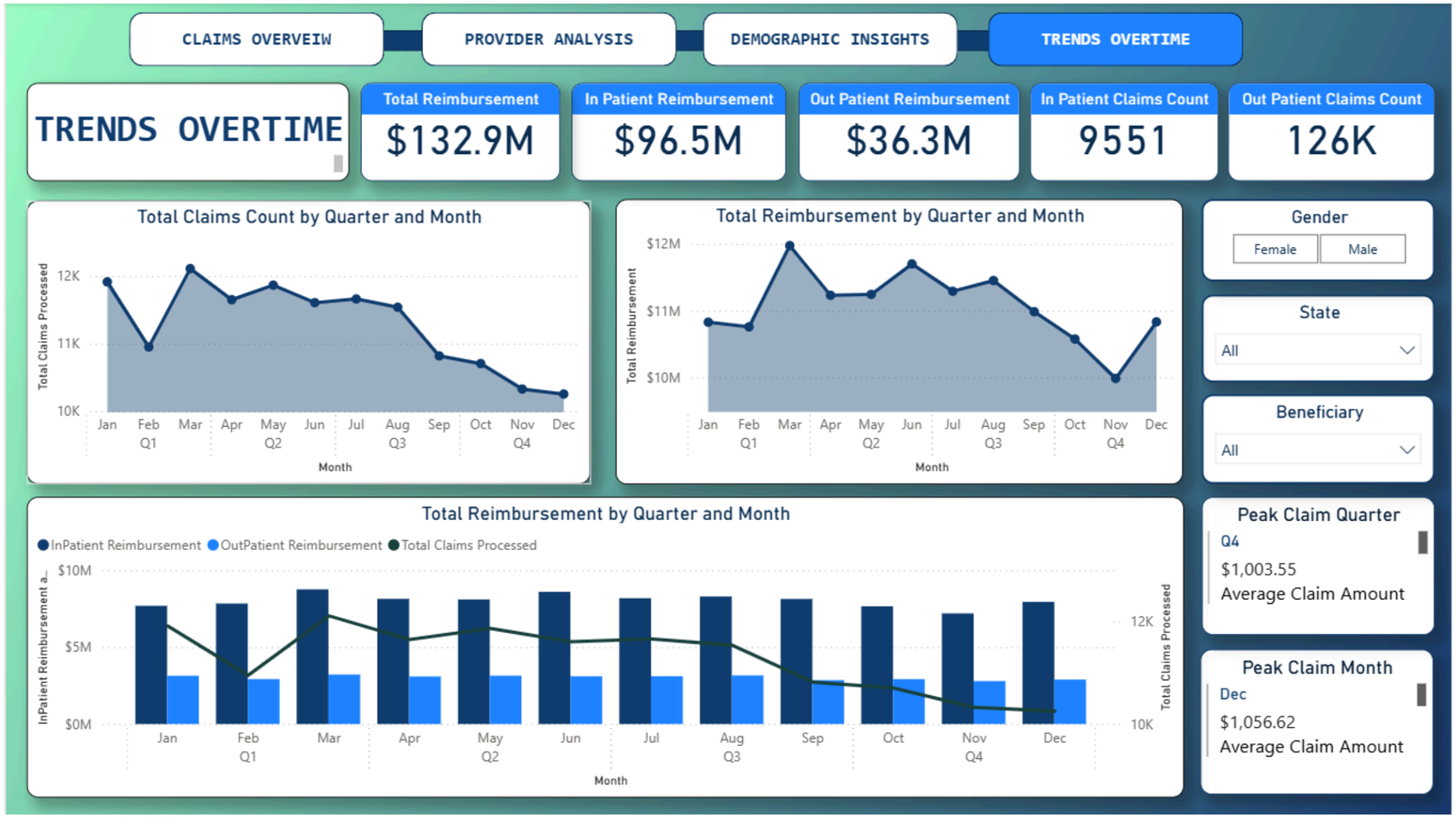
- 57% male, 43% female.
- One race group dominates at 85%, with others much smaller.

But the defining story lies in chronic conditions:

- Diabetes touches 86K lives.
- Ischemic Heart Disease affects 82K.
- Other widespread conditions include: Cancer (119K), Osteoporosis (109K), Kidney Disease (105K), and Depression (103K).

In short: chronic illness shapes the cost and care landscape more than demographics alone.

TRENDS OVERTIME



TRENDS OVERTIME

Claims Journey

- February saw a surge with 12K+ claims.
- Post-September, volumes steadily declined, closing on a slowdown.

Reimbursement Shifts

- February also peaked in value ($12M).
- December flipped the pattern—fewer claims, but the highest cost per claim, signaling costlier treatments.

Inpatient vs. Outpatient

- Inpatients drive the dollars: fewer cases, but $96.5M.
- Outpatients drive the volume: 126K claims, at far lower value per case.

Seasonal Standouts

- Q4 emerges as costliest, averaging $1,003.55 per claim.
- December tops the year: smaller claim counts, but the highest per-claim average ($1,056.62).

THANK YOU
FOR YOUR ATTENTION

Git HUB - https://github.com/Binay005X/Capstone-Project---MediCare-Claims-Hub