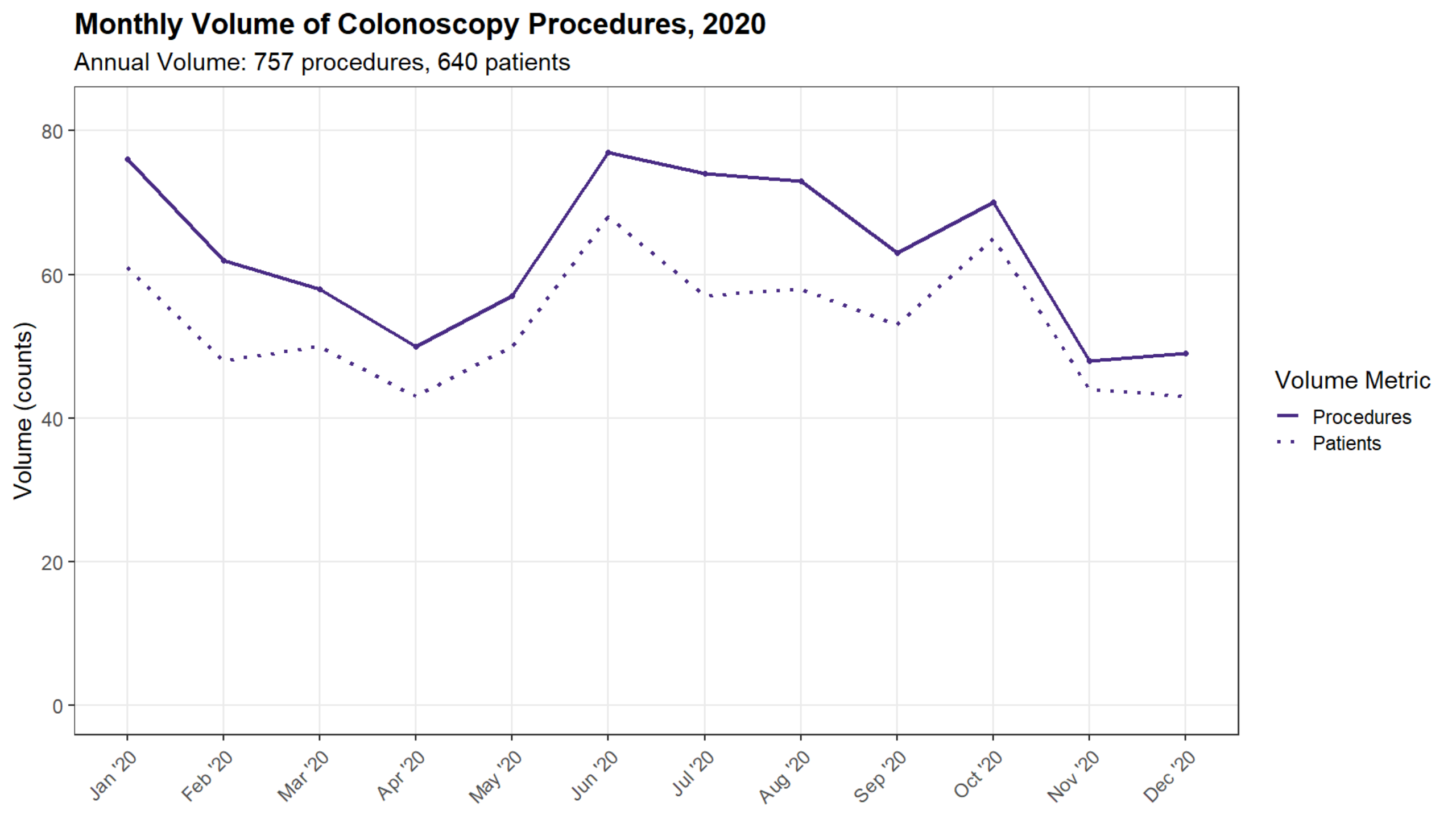
Claims Analysis Report

Georgia Jenkins

Methods: Data Cleaning

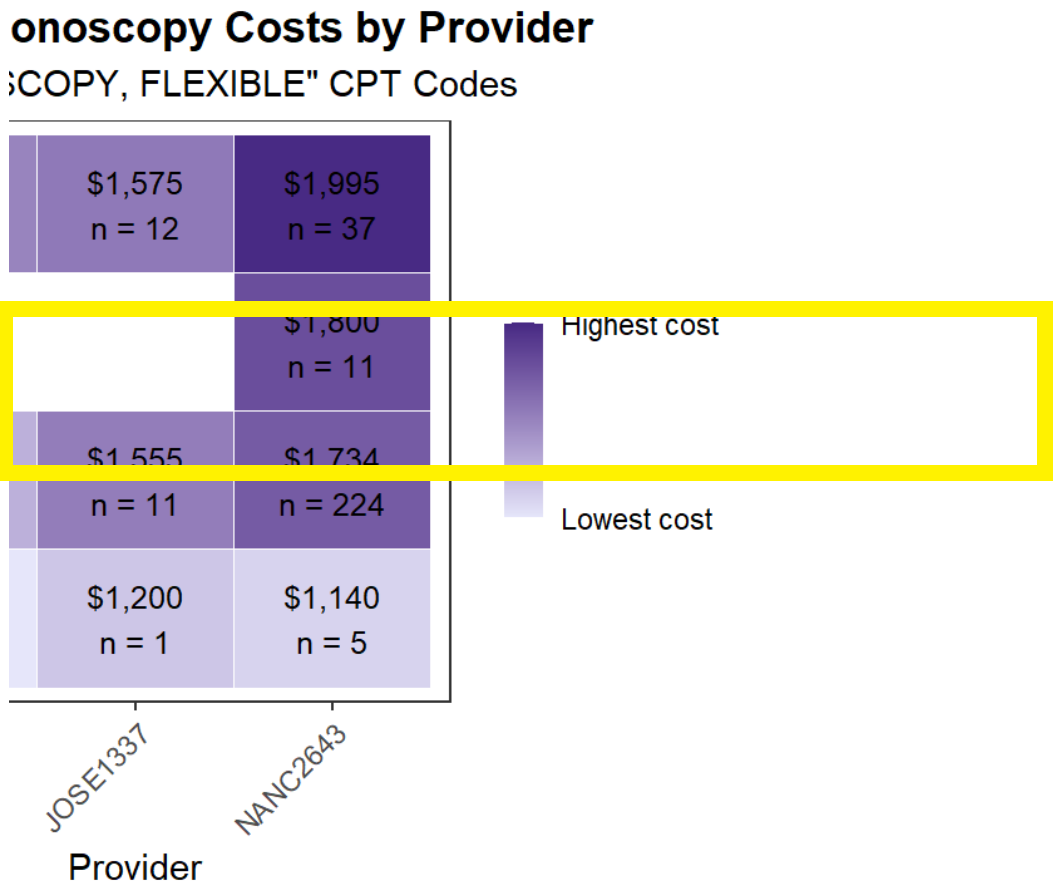
Purpose	Actions Taken
Remove data irrelevant to the assignment	Filtered out endoscopy CPT codes (43200, 43235, 43239). Filtered out procedures that started/ended in any year other than 2020. Confirmed that all provider zip codes were near UHC.
Remove duplicates	Identified claims where claim number, provider, start date, end date, and CPT code were duplicated. 12 duplicate sets (24 claims) were identified. Each set contained one claim with paid amount > 0 and one = 0. Filtered out the 12 duplicates where paid amount = 0.
Remove invalid data	Invalid date, 2/29/2019 was removed during filtering year to 2020. Checked for NA values throughout the data set. None were present in the key variables, just reject codes, modifiers, and additional ICD codes after the primary code. Filtered out 4 claims where payment value = 999999. Assumed this represented an error. Did notice a pattern with the ICD codes, but that didn't adequately explain 999999.


Monthly Volume of Colonoscopy Procedures, 2020
Annual Volume: 757 procedures, 640 patients
Volume (counts)
80
60
40
20
0
Jan '20
Feb '20
Mar '20
Apr '20
May '20
Jun '20
Jul '20
Aug '20
Sep '20
Oct '20
Nov '20
Dec '20
Volume Metric
Procedures
Patients

91% of procedures performed by top 2 providers

Only 1 of the top 3 performed removal by hot biopsy forceps

	Volume		Cost per Procedure	
Provider	**Procedures**	**Patients**	**Mean**	**Median**
HON-2008	357	296	1,341	1,200
NANC2643	277	260	1,761	1,800
JOSE1337	24	13	1,550	1,500
KARL1316	9	6	3,100	3,000
BHAR2019	8	8	1,800	1,800
MICH2059	7	5	1,800	1,800
DONA4738	6	6	1,500	1,500
DANI1341	4	4	1,575	1,500
MARK1336	3	3	1,100	1,200
KENN1336	2	2	1,500	1,500



Cost differences between top 3 providers

Step 1: Determine if there are statistically significant differences

- ANOVA assumptions
  - ✔ Independent observations
  - ✘ Normal distribution
  - ✘ Homogeneity of variances
- Kruskal-Wallis Test **p ≈ 0**

Significant differences exist!

Step 2: which pairs have differences?

- Dunn's test of multiple comparisons, Bonferroni adjustment
  - HON-2008 & JOSE1337 **p < 0.001**
  - HON-2008 & NANC2643 **p ≈ 0**

```
        Kruskal-Wallis rank sum test

data:  paid_amount by provider
Kruskal-Wallis chi-squared = 217.45, df = 2, p-value < 2.2e-16
```

```
                                  Comparison of x by group
                                        (Bonferroni)
Col Mean-|
Row Mean |    HON-2008    JOSE1337
---------+----------------------
JOSE1337 |   -3.449032
         |     0.0008*
         |
NANC2643 |   -14.71050   -2.116606
         |     0.0000*     0.0514

alpha = 0.05
Reject Ho if p <= alpha/2
```

Summary & Recommendations

Provider	Summary
HON-2008	Lowest prices ($p < 0.001$) Highest volume No hot forceps biopsy Had 8% of claims rejected
JOSE 1337	Volume much lower than NANC & HON No hot forceps biopsy No rejected claims
NANC 2863	Highest prices (sig. different from HON-2008 but not from JOSE1337) 2nd highest volume Only provider who performed removals with hot biopsy forceps Had 4% of claims rejected

BEST VALUE
Caveat: high volume may mean longer wait times to get an appointment

Mid-tier value, lower patient volume
May be able to take more patients

Not recommended due to high cost