Hilda Reyes 4AMWork#1

**ABC Pharmacy**

**Summary:**

Overview: ABC Pharmacy transaction data was collected over a period of six months beginning January 1, 2016, ending June 30, 2016.

Business goals: Model high performing pharmacies, consider closing down non and low performing pharmacies.

Data description: Data reflects transactions (sales and returns) associated with active ABC Pharmacies across the country.

Methodological summary: CSV files containing quantitative data were imported for analyses using R. Pharmacy transactions were analyzed on a state and national level.

Results: Non-performing States were identified, these did not report any sales for the period: New Hampshire, Puerto Rico, Vermont, Rhode Island, Virgin Islands, State labeled “Z”. Analysis identified that only 9.66% of pharmacies are reporting sales. Nationwide percentage of merchandise returns is 0.46%.

Recommendations: Consider closing down low and non-performing pharmacies. Explore online accessibility to reach a larger customer population. Merchandise returns will be analyzed if merchandise return percentage becomes significant.

**Geographic Performance:**

Data collected reflects ABC Pharmacy transactions between January 1, 2016 through June 30,2016. Nationwide, ABC Pharmacy holds a total of 1097 pharmacies, of which, 991 do not report sales; this reflects %97.34 of all pharmacies. Only 9.66% of all pharmacies nationwide report sales during this period.

**State Total Pharmacies Pharmacies w/o sales Pharmacies w/ sales**

CT 42 34 8

DE 4 3 1

MA 51 47 4

ME 21 10 11

NH 2 2 0

NJ 288 256 32

NY 390 374 16

PA 196 162 34

PR 97 97 0

RI 3 3 0

VI 1 1 0

VT 1 1 0

Z\* 1 1 0

ABC Pharmacy derives revenue from 106 pharmacies across 7 States: New Jersey, Pennsylvania, Connecticut, Delaware, Massachusetts, New York and Maine; with number of pharmacies ranging between 1 and 34 pharmacies per State. Total sales transactions reported nationwide: 936,680. Total merchandise returns transactions reported nationwide: 4,318. Nationwide rate of merchandise returns is 0.46%.

**State** **Pharmacies** **Sales** **Returns Percentage merchandise returns**

CT 8 68,911 302 0.44%

DE 1 1,034 0 0.00%

MA 4 79,772 486 0.61%

ME 11 43,334 196 0.45%

NJ 32 439,796 1932 0.44%

NY 16 92,846 650 0.70%

PA 34 210,987 752 0.36%

The five highest performing pharmacies are found in New Jersey, Massachusetts and Pennsylvania:

**State Pharmacy Name Sales**

NJ GNP PHARMACY #18 64,281

NJ GNP PHARMACY #795 48,599

NJ GNP PHARMACY #232 40,375

MA GNP PHARMACY #990 37,275

PA GNP PHARMACY #680 36,997

The five lowest performing pharmacies are found in New York, New Jersey and Pennsylvania:

**State Pharmacy Name Sales**

NY GNP PHARMACY #835 121

NY GNP PHARMACY #881 99

NJ GNP PHARMACY #211 27

PA GNP PHARMACY #116 7

PA GNP PHARMACY #150 4

The five pharmacies with highest reported merchandise returns are found in New Jersey, Massachusetts and New York.

**Pharmacy Name State Returns**

GNP PHARMACY #18 NJ 345

GNP PHARMACY #166 NJ 277

GNP PHARMACY #269 NJ 268

GNP PHARMACY #990 MA 216

GNP PHARMACY #977 NY 211

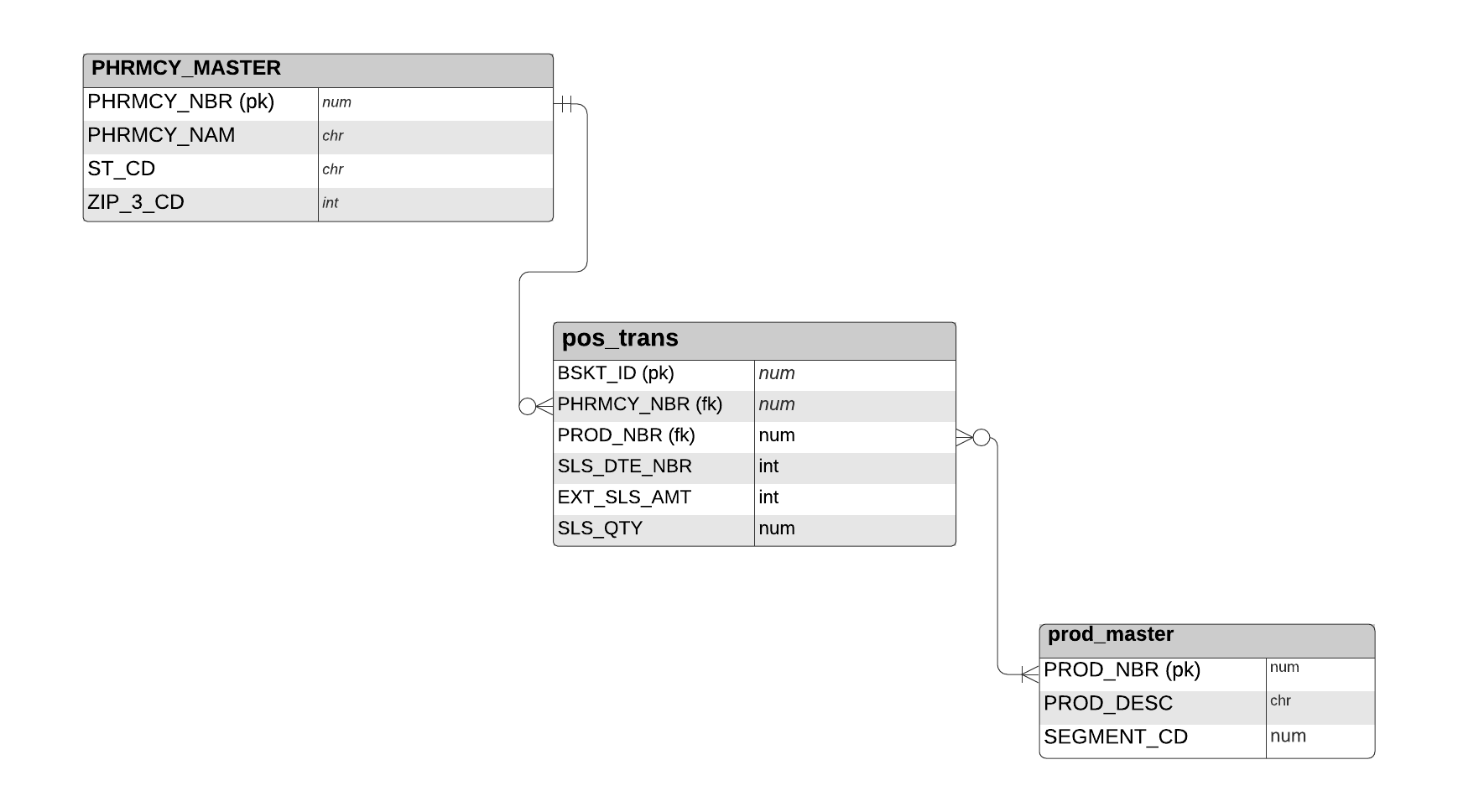
GNP Pharmacy#18 reports the highest number of sales: 64,281, as well as the highest number of merchandise returns: 345. With a merchandise return percentage of 0.54%, GNP Pharmacy#18 is above ABC Pharmacies national rate of merchandise return percentage of 0.46% and New Jersey’s percentage of 0.44%. New York and Massachusetts hold the highest percentage of merchandise returns, 0.70% and 0.61%, respectively. The merchandise return percentage is not significant. Product merchandise returns will be analyzed if merchandise return percentage were significant.

**Data Quality**:

1. State “Z” was identified, but did not impact analyses due to inactivity
2. Calculating Sales required querying (EXT\_SLS\_AMT) transactions > 0
3. Calculating Returns required querying (EXT\_SLS\_AMT) transactions < 0
4. Data frame column name SEGMENT\_CD was changed to SEG\_CD
5. Data frame column type for SEG\_CD was changed from Numeric to Character

**Data Model**

The following tables were used:



1. Data frames PHRMCY\_MASTER and pos\_trans were merged and assigned to the variable Pharm\_trans

Pharm\_trans <- merge(PHRMCY\_MASTER, pos\_trans, by= "PHRMCY\_NBR")

1. Data frames Pharm\_trans and prod\_master were merged and assigned to the variable Pharm\_trans\_by\_prod\_master

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

**Questions, Queries and Results:**

# 1. What is the date range?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% summarise(start\_date=min(SLS\_DTE\_NBR), end\_date=max(SLS\_DTE\_NBR))

## start\_date end\_date

# 1 20160101 20160630

# 2. Which States have pharmacies with reported sales?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% distinct(ST\_CD)%>% arrange(ST\_CD)

## ST\_CD

# 1 CT

# 2 DE

# 3 MA

# 4 ME

# 5 NJ

# 6 NY

# 7 PA

# 3. How many pharmacies nationwide report sales?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR") Pharm\_trans\_by\_prod\_master%>% summarise(Total\_pharm = n\_distinct(PHRMCY\_NAM))

## Total\_pharm

# 1 106

# 4. How many pharmacies per State report sales?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR") Pharm\_trans\_by\_prod\_master%>% group\_by(ST\_CD)%>% summarise(pharm\_count = n\_distinct(PHRMCY\_NAM))

## A tibble: 7 x 2

# ST\_CD pharm\_count

# <chr> <int>

# 1 CT 8

# 2 DE 1

# 3 MA 4

# 4 ME 11

# 5 NJ 32

# 6 NY 16

# 7 PA 34

# 5. What is total sales transactions for all pharmacies nationwide?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% summarise(Total = sum(EXT\_SLS\_AMT>0))

## Total

# 1 936680

# 6. What is total sales transactions per State?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% group\_by(ST\_CD)%>% summarise(Total = sum(EXT\_SLS\_AMT>0))

## ST\_CD Total

# <chr> <int>

# 1 CT 68911

# 2 DE 1034

# 3 MA 79772

# 4 ME 43334

# 5 NJ 439796

# 6 NY 92846

# 7 PA 210987

# 7. What is total transactions of merchandise returns nationwide?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% summarise(Total = sum(EXT\_SLS\_AMT<0))

## Total

# 1 4318

# 8. What is the total transactions of merchandise returns by State?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% group\_by(ST\_CD)%>% summarise(Total = sum(EXT\_SLS\_AMT<0))

## ST\_CD Total

# <chr> <int>

# 1 CT 302

# 2 DE 0

# 3 MA 486

# 4 ME 196

# 5 NJ 1932

# 6 NY 650

# 7 PA 752

#9. What is the Nationwide percentage of returned merchandise?

percent(4318/936680, accuracy = 0.01) Result: # "0.46%"

#10. What is the percentage of returned merchandise per State?

# CT percent(302/68911, accuracy = 0.01) Result: # "0.44%"

# DE percent(0/1034, accuracy = 0.01) Result: # "0.00%"

# MA percent(486/79772, accuracy = 0.01) Result: # "0.61%"

# ME percent(196/43334, accuracy = 0.01) Result: # "0.45%"

# NJ percent(1932/439796, accuracy = 0.01) Result: # "0.44%"

# NY percent(650/92846, accuracy = 0.01) Result: # "0.70%"

# PA percent(752/210987, accuracy = 0.01) Result: # "0.36%"

# 11. Which five stores have the highest merchandise returns?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% group\_by(PHRMCY\_NAM, ST\_CD)%>% summarise(Total = sum(EXT\_SLS\_AMT<0))%>% arrange(desc(Total))%>% head(5)

## A tibble: 5 x 3

# Groups: PHRMCY\_NAM [10]

## PHRMCY\_NAM ST\_CD Total

# <chr> <chr> <int>

# 1 GNP PHARMACY #18 NJ 345

# 2 GNP PHARMACY #166 NJ 277

# 3 GNP PHARMACY #269 NJ 268

# 4 GNP PHARMACY #990 MA 216

# 5 GNP PHARMACY #977 NY 211

# 12. Which are the five highest performing pharmacies?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% group\_by(ST\_CD, PHRMCY\_NAM)%>% summarise(Total = sum(EXT\_SLS\_AMT>0))%>% arrange(desc(Total))%>% head(5)

## # A tibble: 5 x 3

# Groups: ST\_CD [3]

# ST\_CD PHRMCY\_NAM Total

# <chr> <chr> <int>

# 1 NJ GNP PHARMACY #18 64281

# 2 NJ GNP PHARMACY #795 48599

# 3 NJ GNP PHARMACY #232 40375

# 4 MA GNP PHARMACY #990 37275

# 5 PA GNP PHARMACY #680 36997

#13. What is the percentage of returned merchandise for GNP PHARMACY #18?

percent(345/64281, accuracy = 0.01) Result: #[1] "0.54%"

# 14. What are the five lowest performing pharmacies?

Pharm\_trans\_by\_prod\_master<-merge(Pharm\_trans, prod\_master, by="PROD\_NBR")

Pharm\_trans\_by\_prod\_master%>% group\_by(ST\_CD, PHRMCY\_NAM)%>% summarise(Total = sum(EXT\_SLS\_AMT>0))%>% arrange(desc(Total))%>% tail(5)

## A tibble: 5 x 3

# Groups: ST\_CD [3]

# ST\_CD PHRMCY\_NAM Total

# <chr> <chr> <int>

# 1 NY GNP PHARMACY #835 121

# 2 NY GNP PHARMACY #881 99

# 3 NJ GNP PHARMACY #211 27

# 4 PA GNP PHARMACY #116 7

# 5 PA GNP PHARMACY #150 4

#15. What are the total pharmacies per State?

all\_pharms<-PHRMCY\_MASTER%>% left\_join(pos\_trans, by = c("PHRMCY\_NBR" = PHRMCY\_NBR"))%>% group\_by(ST\_CD)%>% summarise(pharm\_count = n\_distinct(PHRMCY\_NAM))

all\_pharms

## A tibble: 13 x 2

# ST\_CD pharm\_count

# <chr> <int>

# 1 CT 42

# 2 DE 4

# 3 MA 51

# 4 ME 21

# 5 NH 2

# 6 NJ 288

# 7 NY 390

# 8 PA 196

# 9 PR 97

# 10 RI 3

# 11 VI 1

# 12 VT 1

# 13 Z 1

#16. What is the total number of pharmacies nationwide?

Total\_Pharm\_nation<-PHRMCY\_MASTER%>% left\_join(pos\_trans, by = c("PHRMCY\_NBR" = "PHRMCY\_NBR"))%>% summarise(pharm\_count = n\_distinct(PHRMCY\_NAM))

Total\_Pharm\_nation

## pharm\_count

# 1 1097

#17. How many pharmacies per State **do not** report sales?

no\_sales<-PHRMCY\_MASTER%>% left\_join(pos\_trans, by = c("PHRMCY\_NBR" = "PHRMCY\_NBR"))%>% filter(is.na(EXT\_SLS\_AMT))%>% group\_by(ST\_CD)%>% summarise(pharm\_count = n\_distinct(PHRMCY\_NAM))

no\_sales

## ST\_CD pharm\_count

# <chr> <int>

# 1 CT 34

# 2 DE 3

# 3 MA 47

# 4 ME 10

# 5 NH 2

# 6 NJ 256

# 7 NY 374

# 8 PA 162

# 9 PR 97

# 10 RI 3

# 11 VI 1

# 12 VT 1

# 13 Z 1

#18. How many pharmacies nationwide **do not** report sales?

Total\_Pharm\_no\_sales<-PHRMCY\_MASTER%>% left\_join(pos\_trans, by = c("PHRMCY\_NBR" = "PHRMCY\_NBR"))%>% filter(is.na(EXT\_SLS\_AMT))%>% summarise(pharm\_count = n\_distinct(PHRMCY\_NAM))

Total\_Pharm\_no\_sales

## pharm\_count

# 1 991

#19. What is the percent of pharmacies nationwide that **do not** report sales?

percent(991/1097, accuracy = 0.01) Result: # "90.34%"

#20. What is the percent of pharmacies nationwide reporting report sales?

percent(106/1097, accuracy = 0.01) Result: # "9.66%"

Top of Form