THE STATE UNIVERSITY OF NEW JERSEY

RUTGERS

# Visual Audit Task

Qi Liu, Lu Zhang, and Heejae Lee

# Agenda

- Introduction
- Literature Review
- Data
- Visual Audit Exploratory Data Analysis Framework
- Visual Audit Task 1: Revenue Analysis
- Visual Audit Task 2: Internal Control Test
- Visual Audit Task 3: Testing Occurrence Assertion
- Visual Audit Task 4: Transaction Value Analysis
- Visual Audit Task 5: Analytical Procedure
- Questions
- Future Work

### Introduction

- Using interactive visualization in accounting information system can help users to understand a very large set of financial information.
- In particular, continuous monitoring and continuous auditing increases the value of interactive visualization in an accounting information system context.
- However, there are only few studies that examine how interactive visualization should be applied in accounting, especially in auditing.
- Examine how interactive visualization can be used in auditing and suggest novel framework for visual audit.
- Demonstrate five visual audit task using hospital database

### Literature Review

- Prior literature examined when and how interactive data visualization can bring more effective and efficient decision making for accounting data (Dilla & Raschke, 2010, Yigitbasioglu and Velcu, 2012).
- Most researches suggested that the effectiveness of data visualization tool depends on task characteristics and decision makers characteristics, consistent to cognitive fit theory.
- Dilla and Raschke (2015) suggested a framework for interactive data visualization in fraud detection procedure.
- However, the frameworks suggested in previous studies are theoretical, rather than practical.
- In this study, we applied explanatory data analysis (EDA) approach to develop visual audit framework.

### Data

- Data from 45 hospitals
- Select client 11 as a target and perform visual audit task
  - Midsize hospital (Transaction volume, Number of Accounts)
  - High proportion of Bad Debt
- Use transactions from 2014 to 2016 for the analysis
  - Data integrity issue for previous year data
- Total 6,093,847 transactions record for client 11

## Data

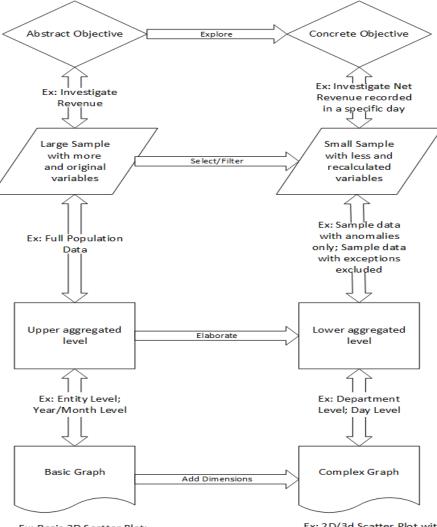
- Tables
  - Accounts
    - Shows most recent information for an account.
  - AR Snapshot
    - Created from ATB (aged trial balance) files
    - Shows account balances, aging, and general information for an account within a hospital system at end-of-month
  - AR Transaction
    - Shows the transactions that occurred, by posting date, for accounts within given hospital systems
  - AR Snapshot Estimates
  - MRA Snapshot Estimates
    - Holds the estimates that were used to create reserves by account during a hospital systems monthly close process
  - VA Gross Revenue
    - Shows the gross and net revenue attributed to an account as of a period id.

## Data

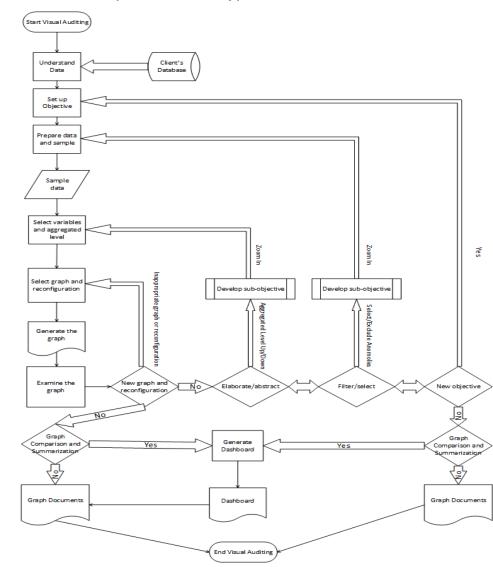
- Mapping Tables
  - Billing Type, Financial Class, Inpatient/Outpatient Type, Insurance Provider, Patient Type, etc.
  - Service Type
    - Inpatient, Outpatient, Ambulatory Surgery, Emergency (ER), PCU (Progressive Care unit), Intensive Care Unit (ICU), Private, etc.
  - Transaction Type
    - Administration Adjustment, Bad Debt, Charity, Contractual, Denial, Exclude, Not Specified, and Payment

### Visual Audit using Exploratory Data Analysis Framework

Upper level (Zoom in Approach)

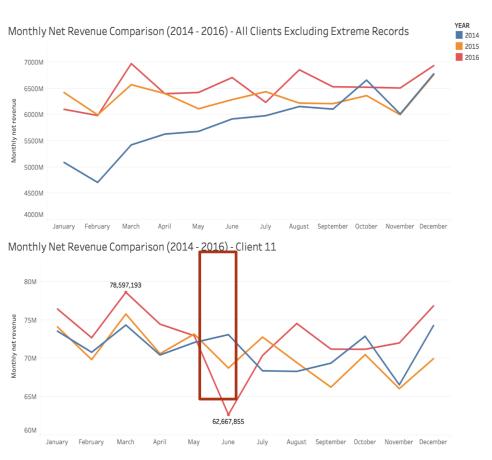


Ex: Basic 2D Scatter Plot; Bar Chart Ex: 2D/3d Scatter Plot with color, size, shape and transparence difference; Stacked/Clustered Bar Chart Lower level (Visual EDA Loop)



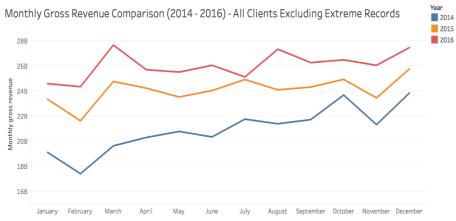
### Visual Audit Task 1: Revenue Analysis

- Mainly perform risk assessment
- Demonstrate the monthly net revenue, gross revenue, net revenue over gross ratio and account receivable over net revenue ratio
- Drill down from entity level to lower level (e.g. account level)

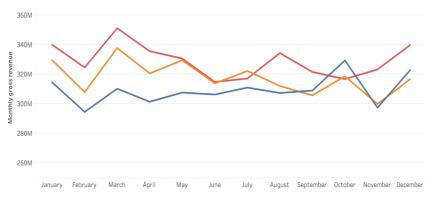


RUTGERS

#### Dashboard 1: Total Net Revenue Monthly Analysis



Monthly Gross Revenue Comparison (2014 - 2016) - Client 11



#### Dashboard 2: Total Gross Revenue Monthly Analysis

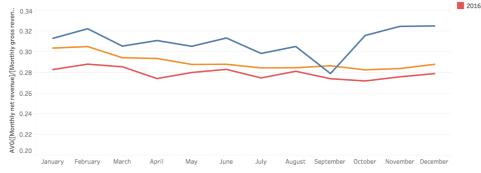
Year

2014

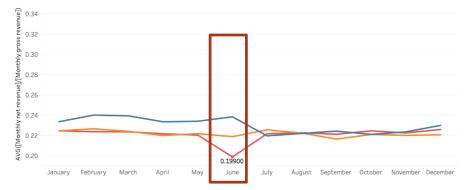
2015

Monthly Net Revenue over Gross Ratio Comparison (2014 - 2016) - All Clients Excluding Extreme Records

RUTGERS







Dashboard 3: Net Revenue over Gross Revenue Ratio analysis

We found that

- Net revenue over gross revenue ratio is lower than the average ratio of all clients', which might indicate higher risk.
- The monthly changes of ratios of client 11 are smaller than all clients'
- A big drop on June 2016 for client 11, probably caused by the big net revenue drop in that month.

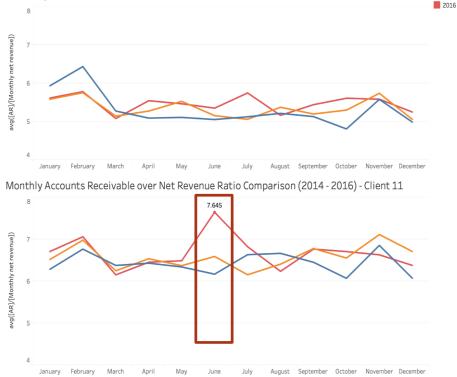
Year

2014

2015

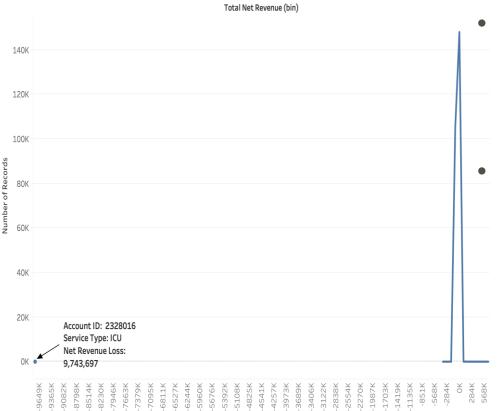
Monthly Accounts Receivable over Net Revenue Ratio Comparison (2014 - 2016) - All Clients Excluding Extreme Records

RUTGERS

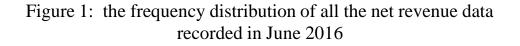


Dashboard 4: Accounts Receivable Net Revenue Ratio analysis

- Use Account Balance data recorded by the hospitals database as the Accounts Receivable
- We found that
  - The client's Accounts Receivable over Net Revenue ratio is higher than the average ratio of all clients, which may indicate higher risk.
  - A big ratio increase on June 2016, which may be due to the big revenue drop.

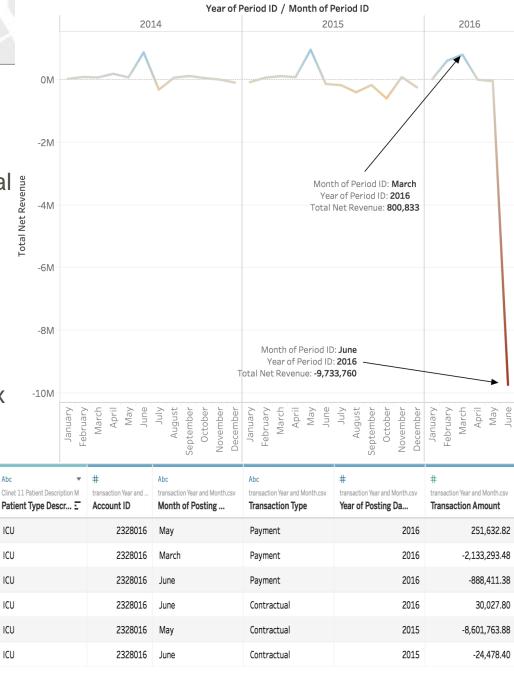


TGERS



- Drill down to the account
  level to investigate the
  reason for the revenue drop in June 2016.
- Create line chart of the frequency distribution of all net revenue data recorded in June 2016 and find one extreme exception, account 2328016
  - The exceptional record of net revenue loss might be the main cause for the big revenue drop in June 2016.

- Net revenue loss recorded in June 2016 was exceptional and abnormal
- No further revenue reports after June 2016, which indicated that the hospital may close the account 2328016 after June 2016
- The account is recorded under the ICU service type which only contributes about 4% revenue to the total revenue.
- In transaction table, we only found six months from 2014 to 2015 that recorded transaction history.
- However, in the revenue table, we found the complete revenue records from January 2014 to June 2016.
- Further testing on Accuracy and Occurrence assertions might be necessary.



Dashboard 5. Line chart showing the monthly net revenue gain/loss reported under account 2328016. Table shows the history of transaction of the account.

RUTGERS

- Table Accounts shows the aggregated transaction values and the most recent balances for each account.
- Table Transaction shows the daily transactions under different transaction types for each account.
- In theory, the calculated total transaction value for each accounts using data from Table – Transaction should equal to the total transaction value reported for the same account in the Table – Accounts.
- A large deviation between two values might indicate internal control deficiency.

RUTGERS



Dashboard 16: The scatter plot show total transaction value in Account Table (X axis) and sum of transaction value of the account in Transaction Table (Y axis)

• Scatterplot analysis shows that

RUTGERS

- There are some accounts with mismatched recorded values
- There are more accounts with positive difference than those with negative difference.
- Larger number of accounts with mismatched bad debts records than that of other transaction types.

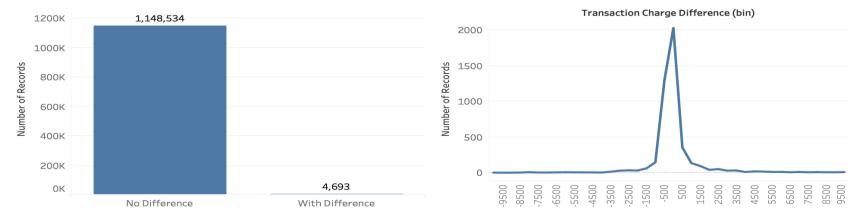
RUTGERS

- Total charge: the value hospital charged the customer
- Total transaction values: charged balances were reduced by payment, admin adjustment, bad debts, contractual and charity deduction.
- Therefore, for each closed account, the total charge value was supposed to be equaled to the final total transaction value reported in the Table Account.
- Value differences may indicate possible internal control weakness or completeness and accuracy assertion violation.
- Take a full population of accounts opened after January 2014 and closed before December 2016

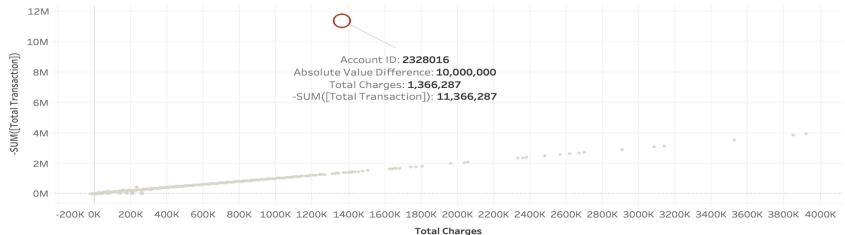
Number of Accounts with No Value Difference versus with Value Difference

RUTGERS

Value Difference Distribution (Absolute Value More Than 1 and less than 10K)



#### Total Transaction Value versus Total Charges



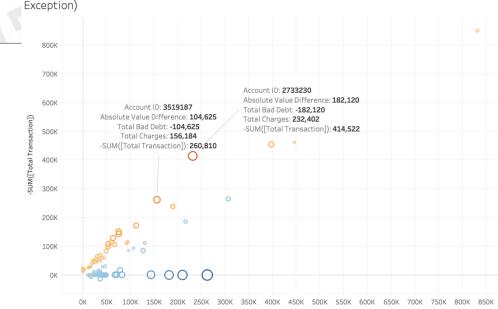
Dashboard 17

#### Rutgers

## Visual Audit Task 2: Internal Control Test

- There was one account: 2328016 with 10M value difference, which may be caused by error
- Only about 0.4% closed accounts reported value difference
- There was no abnormal hypes or patterns detected regarding the frequency distribution of accounts with value difference.

- Exclude Account 2328016, which might be a outlier or exception
- Identify 134 accounts with more than 10K value difference
- Later, auditor could further investigate those accounts if there was any internal control weakness, error or fraud causing those large value differences



Total Transaction Value versus Total Charges (Difference more than 10K and Excluding

Total Charges

Total Exclud	Total Contractual	Total Charity	Total Charges	Total Bad Debt	Total Admin Adj	ABS([Total Charges]+[Total Transaction])	([Total Charges]+[Total Transaction])	-SUM([Total Transaction])	Account ID
0.0	0.00	0.00	182,368.63	0.00	0.00	182,368.63	182,368.63	0.00	3954463
0.0	-13,201.01	0.00	23,213.77	0.00	0.00	10,012.76	10,012.76	13,201.01	3946816
0.0	0.00	0.00	40,106.36	0.00	0.00	40,106.36	40,106.36	0.00	3921628
0.0	0.00	0.00	10,933.82	0.00	0.00	10,933.82	10,933.82	0.00	3899153
0.0	-84,566.63	0.00	127,628.42	0.00	0.00	43,061.79	43,061.79	84,566.63	3879838
0.0	-9,954.98	0.00	33,172.98	0.00	0.00	19,215.82	19,215.82	13,957.16	3874189
0.0	0.00	0.00	16,690.63	0.00	0.00	16,690.63	16,690.63	0.00	3870373
0.0	-31,701.35	0.00	27,873.23	0.00	0.00	17,916.83	-17,916.83	45,790.06	3868597
0.0	0.00	0.00	0.00	0.00	0.00	11,615.63	-11,615.63	11,615.63	3864035
0.0	-35,566.15	0.00	45,566.15	0.00	0.00	12,171.01	-12,171.01	57,737.16	3845592
0.0	1,273.94	0.00	26,748.57	0.00	0.00	26,748.57	26,748.57	0.00	3820400
0.0	0.00	0.00	17,359.52	0.00	0.00	17,359.52	17,359.52	0.00	3796253
0.0	0.00	0.00	12,630.00	0.00	0.00	12,630.00	12,630.00	0.00	3794853
0.0	-29,119.26	0.00	47,506.04	0.00	0.00	18,386.78	18,386.78	29,119.26	3784286
0.0	0.00	0.00	13,384.51	0.00	0.00	13,384.51	13,384.51	0.00	3773906
0.0	-77,957.16	0.00	92,485.95	0.00	0.00	17,484.20	-17,484.20	109,970.15	3752959
0.0	0.00	0.00	12,148.00	0.00	0.00	12,148.00	12,148.00	0.00	3726805
0.0	-46,762.84	0.00	107,689.57	0.00	0.00	15,637.00	15,637.00	92,052.57	3724261
0.0	-29,651.61	0.00	46,694.89	0.00	0.00	17,043.28	17,043.28	29,651.61	3722873
0.0	-7,683.47	0.00	36,365.49	25,034.67	0.00	50,069.34	50,069.34	-13,703.85	3696702
0.0	-1,288.00	0.00	15,648.20	0.00	0.00	13,885.00	13,885.00	1,763.20	3695818
0.0	-17,073.45	0.00	78,293.83	0.00	0.00	61,220.38	61,220.38	17,073.45	3673362
0.0	0.00	0.00	15,120.06	0.00	0.00	15,120.06	15,120.06	0.00	3640569
0.0	-185,270.41	0.00	216,375.55	0.00	0.00	30,820.14	30,820.14	185,555.41	3640149
0.0	0.00	0.00	11,329.39	0.00	0.00	10,786.97	10,786.97	542.42	3635640
0.0	0.00	0.00	14,472.92	0.00	0.00	14,472.92	14,472.92	0.00	8620154
0.0	0.00	0.00	11,840.33	-23,680.66	0.00	11,840.33	-11,840.33	23,680.66	3593965
0.0	0.00	0.00	33,998.22	0.00	0.00	33,998.22	33,998.22	0.00	3590062
0.0	-54,651.88	0.00	31,421.31	0.00	0.00	32,708.18	-32,708.18	64,129.49	3577102
0.0	0.00	0.00	48,980.13	0.00	0.00	48,980.13	48,980.13	0.00	3573033
0.0	-40,272.99	0.00	97,014.53	0.00	0.00	12,114.94	12,114.94	84,899.59	3530252
0.0	-68,943.44	0.00	156,184.47	-104,625.38	0.00	104,625.27	-104,625.27	260,809.74	3519187
0.0	-18,722.17	0.00	3,714.00	0.00	0.00	15,008.17	-15,008.17	18,722.17	3515898
0.0	0.00	0.00	307,047.02	-222,949.80	0.00	42,048.61	42,048.61	264,998.41	3509143
0.0	-101,497.83	0.00	131,061.29	0.00	0.01	19,563.47	19,563.47	111,497.82	3493506
0.0	-15,298.85	0.00	13,229.36	0.00	0.00	11,530.36	-11,530.36	24,759.72	3491389
0.0	-24,601.62	0.00	37,863.49	0.00	0.00	10,118.23	-10,118.23	47,981.72	8479502
0.0	0.00	0.00	14,602.69	0.00	0.00	14,602.69	14,602.69	0.00	3454076

134 row:

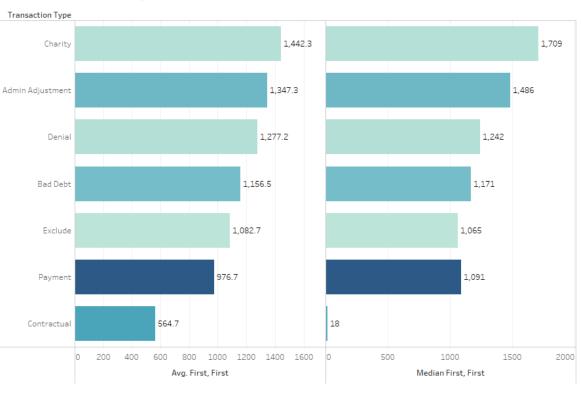
#### Visual Audit Task 3: Testing Occurrence Assertion

- Test Occurrence Assertion using the date difference between billing date and collecting date
- Focus on the difference between first billing date and first transaction recording date
- Change the bin size for the detailed investigation

#### **Visual Audit Task 3: Testing Occurrence Assertion**

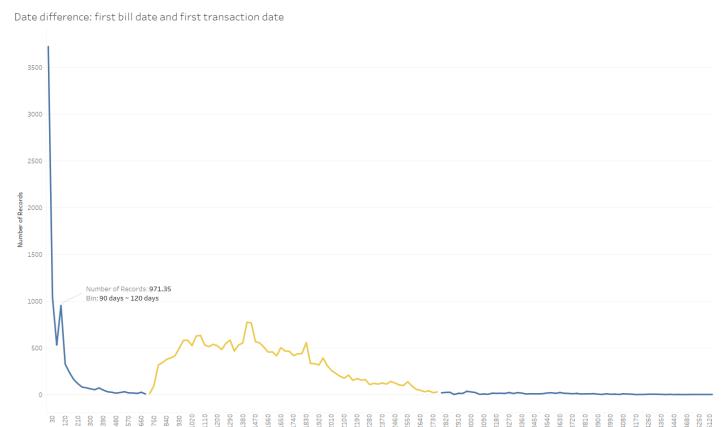
- 1. Date difference between first billing date and first transaction recording date under 7 different transaction types.
- Bar length: Average date difference (left graph) and Median date difference (right graph); Color intensity: Number of records
- Testing sample (about 70k) is small compared to the full population (4M) because most of the billing date is recorded as 19000101.
- 4. Findings: 1) Payment transaction reported the highest frequency, 2) The payment took place 2.5 years after the first billing date on average,3) for the contractual transaction, there is a big difference between average date difference and median date difference

Date Difference: Average and Median



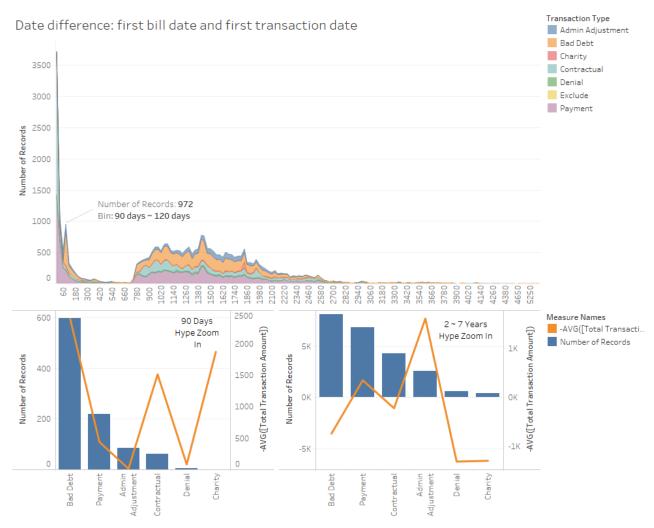
#### **Visual Audit Task 3: Testing Occurrence Assertion**

- Value frequency line chart (bin size = 30 Days)
- Two hypes was identified. The first one occurred at 90 ~ 120 days and the second one occurred around 2.5 ~ 7.5 years.



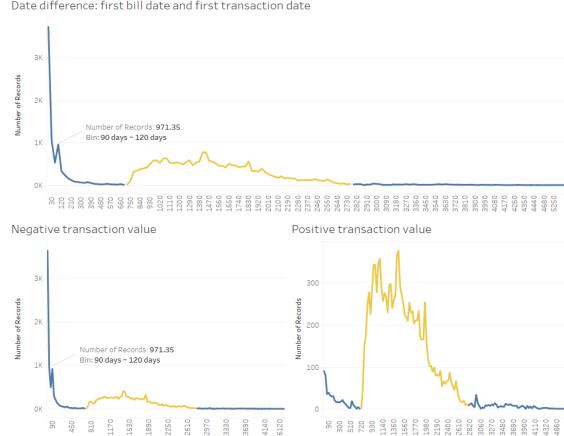
#### Visual Audit Task 3: Testing Occurrence Assertion

- 1. Lower the aggregated level to service level (top figure) and select a sub-sample including accounts only recorded in the first hype (bottom left figure) and the second hype (bottom right figure).
- 2. Findings: 1) both hypes were mainly caused by the bad debt transactions; 2) the second hype was highly related to the transactions with positive numbers (probable adjustment/transfer); 3) the first hype was solely caused by the regular transactions with negative numbers



#### **Visual Audit Task 3: Testing Occurrence Assertion**

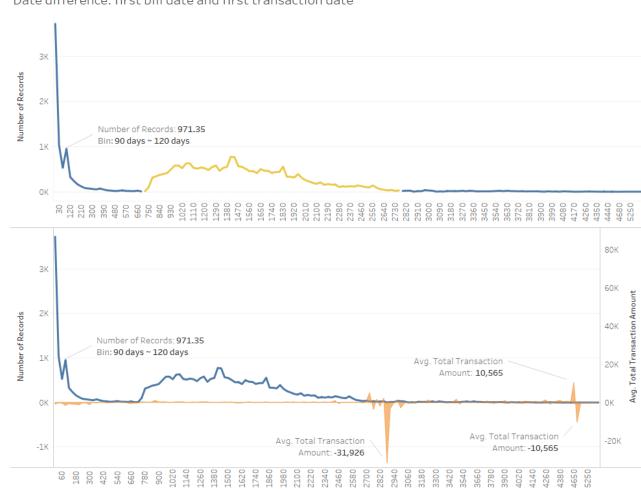
Confirm the second finding stated in the previous slide by generating the similar line chart but using negative value transactions (bottom left figure) or positive value transactions (bottom right figure).



Date difference: first bill date and first transaction date

#### Visual Audit Task 3: Testing Occurrence Assertion

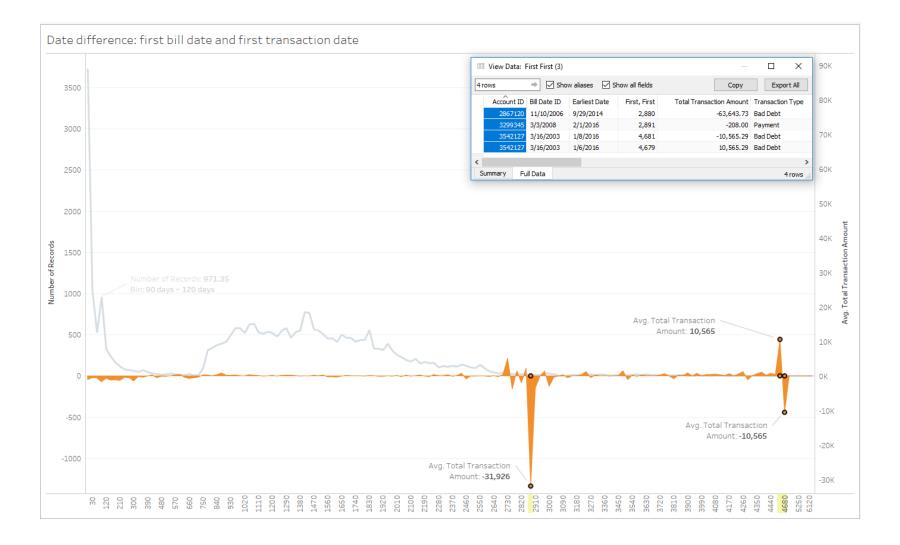
- Use dual axis and orange area to indicate the average recorded transaction value for each bin.
- 2. Anomalies identified (low frequency but high value; offsetting records with high value).



Date difference: first bill date and first transaction date

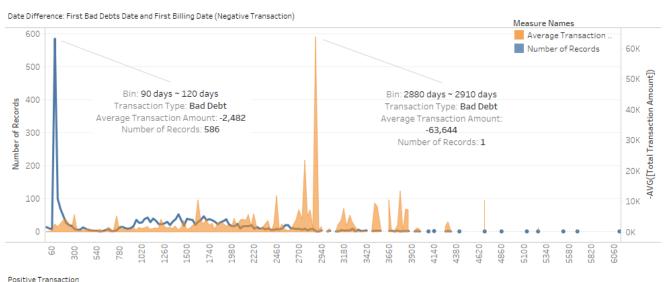
#### **Visual Audit Task 3: Testing Occurrence Assertion**

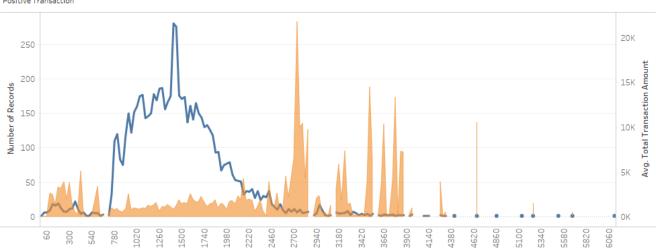
Select anomalies and use table for further investigation.



#### Visual Audit Task 3: Testing Occurrence Assertion

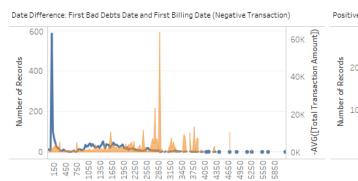
- 1. Similar graph as that of last slide
- 2. Used bad debts transaction records only and divided the sample into transaction with negative values (top figure) and positive values (bottom figure).

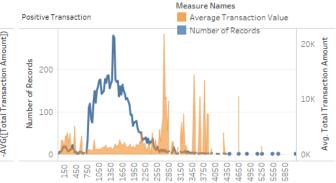




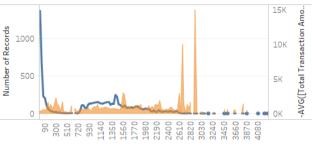
#### Visual Audit Task 3: Testing Occurrence Assertion

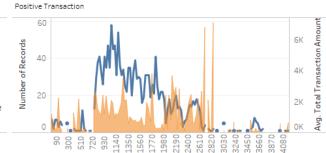
Applied the similar method to other transaction types: Payment (middle figure) and Contractual (bottom figure).







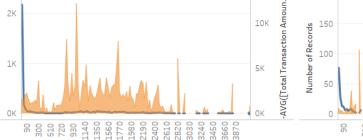


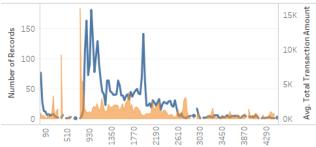




Number of Records







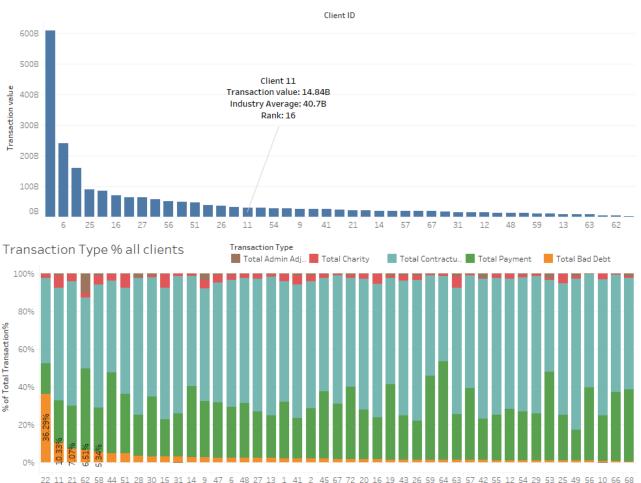
#### Visual Audit Task 4: Transaction Value Analysis

- Perform risk assessment and test audit assertions
- Use transaction values under different Transaction Type
- Conduct K-means clustering analysis to identify comparable peers

#### Visual Audit Task 4: Transaction Value Analysis

1. We generated bar chart to compare total transaction value among all clients (top figure). Based on the results, Client 11 ranked No.16 over 45 hospitals and its reported transaction value was below the industry average (14.84b versus 40.7b).

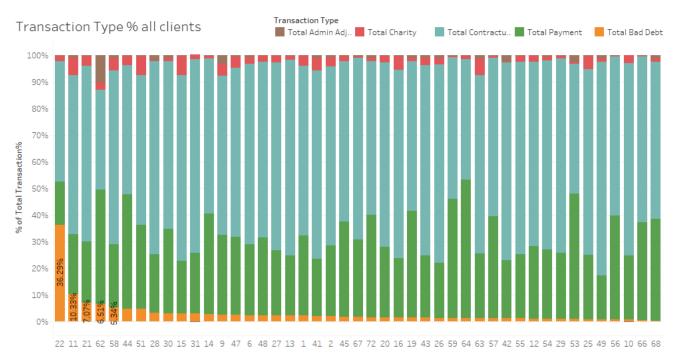
2. Zooming into the transaction type level, we generated stacked bar chart to investigate the percentage of transaction value of different transaction types over total among all clients. We found that client 11 reported high bad debts transaction % ( ranked the second).



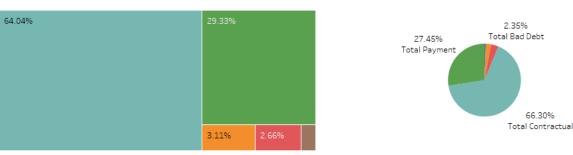
Total Transaction Value all clients

#### **Visual Audit Task 4: Transaction Value Analysis**

We further compared client 11's bad debts percentage with industry average (bottom left) and with the percentage of total bad debts over total transaction value reported by all clients. The bad debts percentage of client 11 was way over those two indicators (10.33% versus 3.11% and 2.35%).



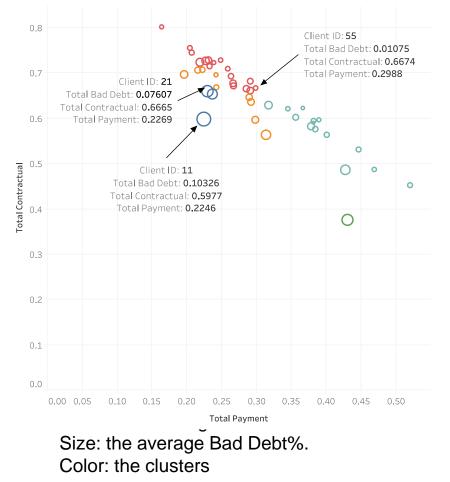
Industry Average



#### Total Bad Debts over Total Transaction - All Clients

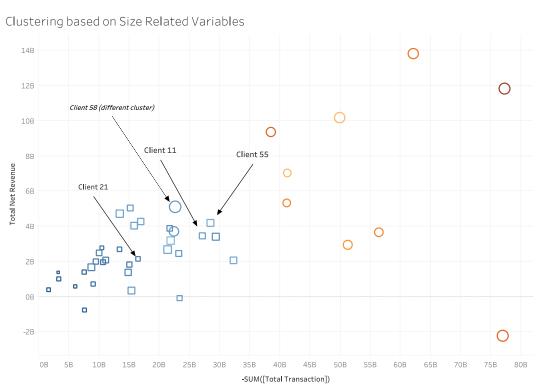
#### Visual Audit Task 4: Transaction Value Analysis

Clusting based on Transaction Type value proportion, excluding Client 22



- Run K-means clustering using transaction structure based on transaction type (Admin Adjustment, Bad Debt, Charity, Contractual, and Payment)
  - Proportion over total transaction value
- Exclude client 22 from the sample because it has exceptionally high bad debt%.
- Client 11 has higher bad debts% and lower payment%
- Drill down from entity level to the division level by analyzing aggregated value based on the service type the client offered.

#### Visual Audit Task 4: Transaction Value Analysis



TGERS

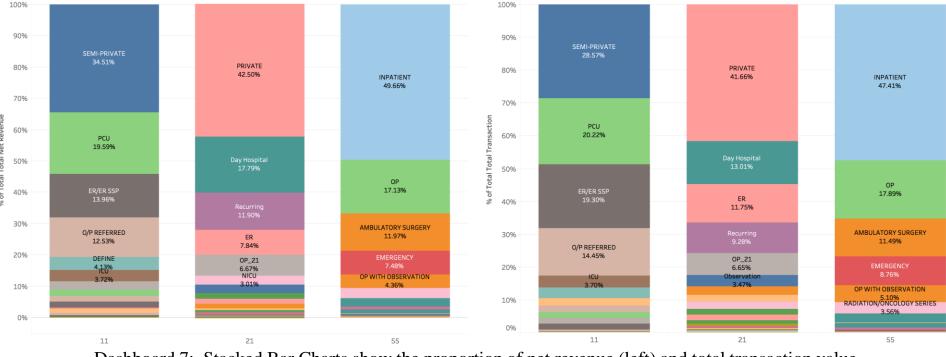
- Run K-means clustering analysis using five size related variables (total account number, total transaction value, total payment value, total net revenue and total account receivable balances)
- Exclude client 1, 2, 6 whose sizes are significantly larger than others.
- Client 55 is similar to Client
   11 in terms of size

#### Visual Audit Task 4: Transaction Value Analysis

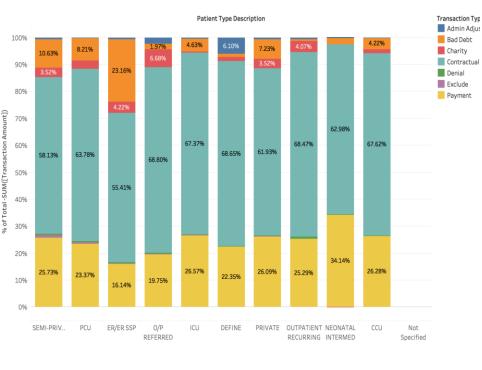
- Select two hospitals as the peers of client 11
- Client 21
  - both belong to the same cluster representing the highest bad debt transaction value proportion over total transaction value and also belong to the same cluster indicating smaller size
- Client 55
  - similar in terms of size to client 11 while it belongs to the cluster with the lowest bad debt transaction over total transaction value.
- Create dashboard to compare three clients' total transaction value and total net revenue under different service types

RUTGERS

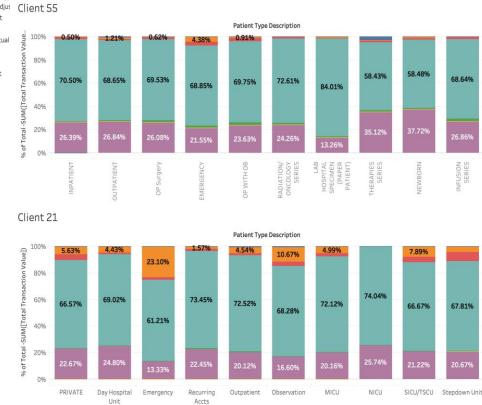
- For client 11 and 21, most of their revenue and transaction values are generated from more special service like semi-private, private, and PCU.
- On the other hand, for client 55, most of its revenue and transaction are carried out by common service such as IP and OP service.



Dashboard 7: Stacked Bar Charts show the proportion of net revenue (left) and total transaction value under service type of each clients (Client 11, Client 21, and Client 55)



RUTGERS

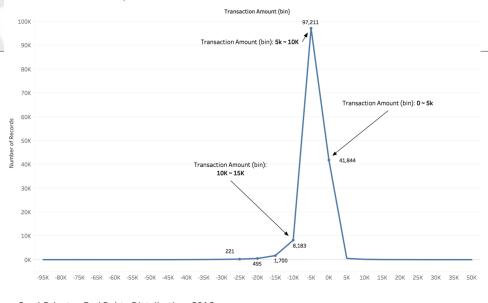


Dashboard 8: Stacked Bar Charts show the proportion of transaction value of different transaction type under Top 10 service type of each clients (Client 11, Client 55, and Client 21).

This dashboard shows that Client 11 has high bad debts%.

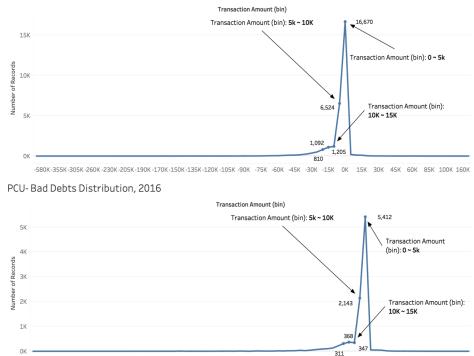
- Drilled down to the account level to investigate why the client has high bad debts % reported under ER, Semi-Private and PCU services.
- We found that
  - Semi-Private and PCU services have more exceptional records with extreme values
  - Semi-Private and PCU services have a lot small but positive transaction records
  - ER has a large number of small to medium bad debts transaction reports

Dashboard 10: Line charts showing the distribution of bad debts transaction of ER, Semi-Private, and PCU



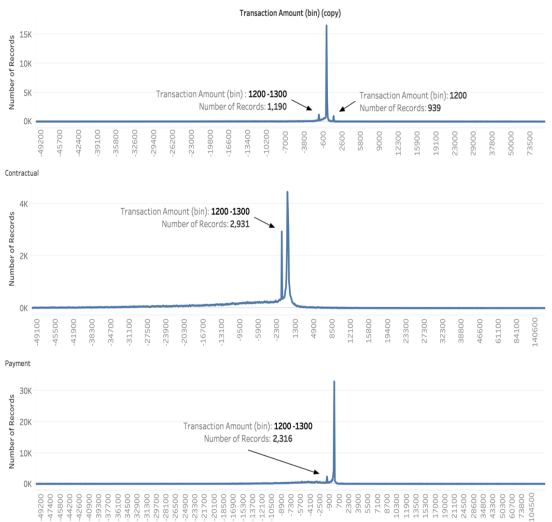
Semi-Private - Bad Debts Distribution, 2016

ER - Bad Debts distribution, 2016



-730K-445K-380K-340K-295K-270K-240K-220K-190K-170K-155K-135K-120K-105K -90K -75K -60K -45K -30K -15K 0K 15K 30K 45K 60K 100K 135K

- Sub-Task: Other transaction types analysis
- Narrow down the bin size from 5000 to 100.
- An abnormal hype located in the bin: \$-1200 \$-1300 for Payment, Contractual and Bad Debts transactions
   and another hype located in Payment the bin: \$1200 \$ 1300 for Bad Debts transactions.
- Next, we narrow down the bin size from 100 to 1 for sub-sample



Dashboard 11: Line charts showing the distribution of Bad Debts, Contractual, and Payment transaction under Semi-Private, PCU, ICU and CCU services

Bad Debts

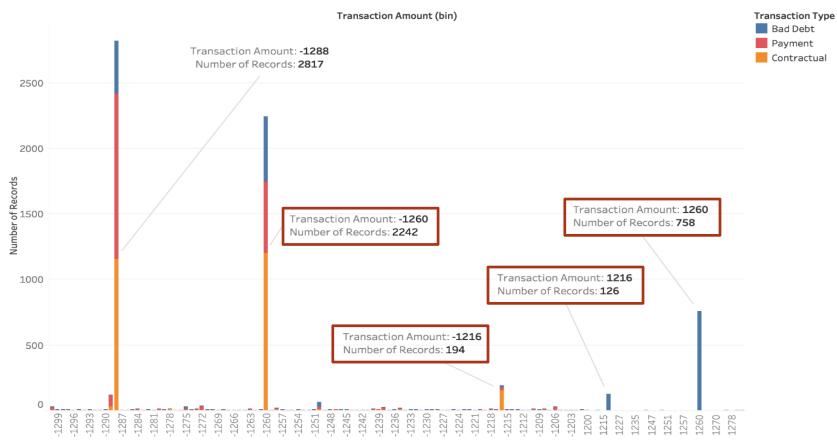
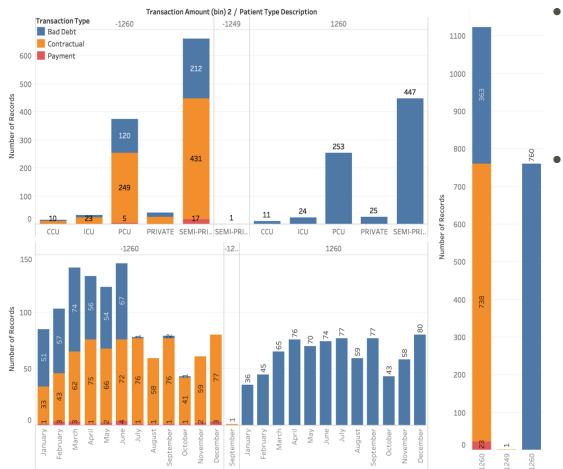


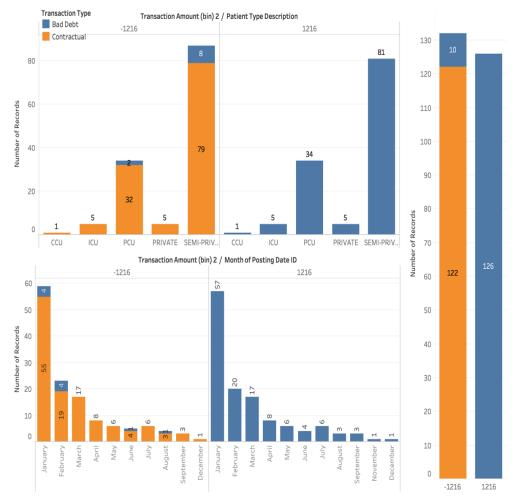
Figure 6: Stacked bar chart show the frequency distribution of the transaction value reported by the accounts



RUTGERS

Dashboard 12: Stacked bar chart show the number of records under different transaction type of bin \$1260 and bin\$-1260 in service level (top) and month level (bottom)

- Drill down the analysis from entity level to service level and from year level to month level.
- We found that
  - For each month, the number of accounts reported bad debts transaction of \$1260 equals to a combined number of accounts reported contractual and payment transaction of \$-1260
  - most accounts fall into those two bins were recorded under Semi-Private and PCU services.



RUTGERS

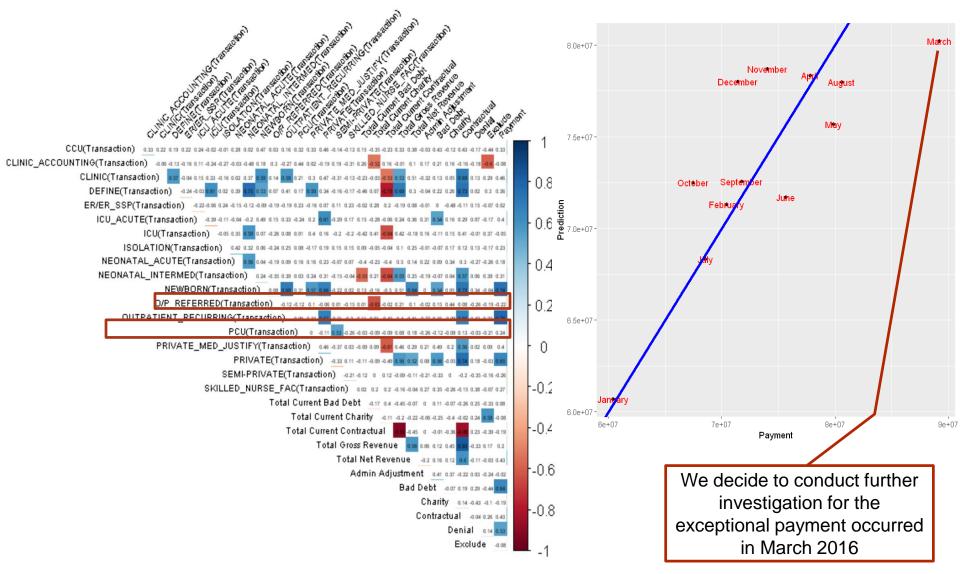
- We found similar results for bin \$1,216 and \$-1,216.
- The client might transferred multiple bad debts balance with same dollar value as either \$1260 or \$1216 to contractual deduction or a combination of contractual and payment transactions.

Dashboard 13: Stacked bar chart show the number of records under different transaction type of bin \$1216 and bin\$-1216 in service level (top) and month level (bottom)

# Visual Audit Task 5: Analytical Procedure

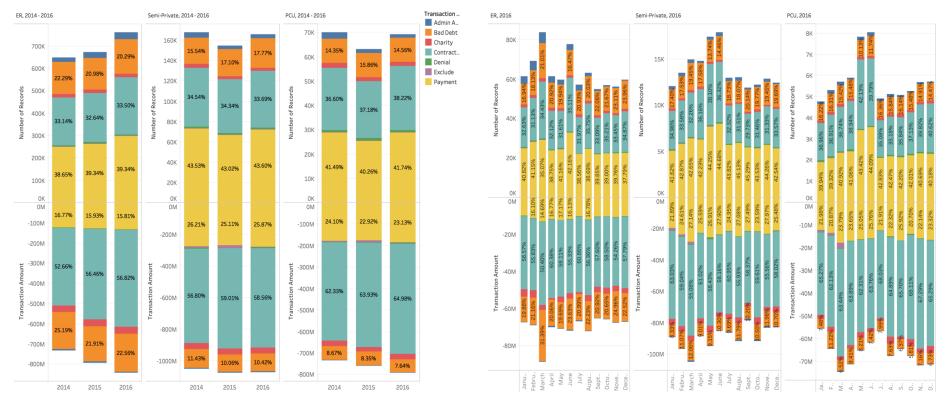
- Investigate the correlations among monthly transaction values under different service types
- Build the prediction model for monthly net revenue, monthly gross revenue and monthly transaction values under different transaction types.

RUTGERS



RUTGERS

 Drill down from entity level to service level (ER, Semi-Private, PCU) and yearly to monthly level



Dashboard 14: Stacked bar chart show the proportion of transaction amount and number of record under different transaction type in different service level, yearly (left) and monthly (right)

• The proportions are constant over three years

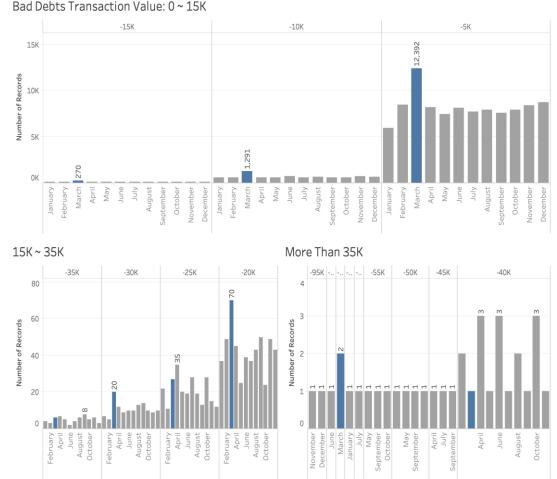
RUTGERS

- Large number of records and transaction value under ER were reported in March 2016
- Large number of records under ER was reported in June 2016 but total transaction value was normal
- Both Semi-Private and PCU reported high transaction frequency in May and June in 2016 but recorded transaction value was normal
- High bad debts transaction value percentage in March 2016
   reported under ER service

 Drill down to the account level and create frequency distribution of Bad Debts bar charts with bin size equal to 5000.

RUTGERS

 High bad debts transaction value reported under ER in March 2016 is not due to exceptional accounts with exceptional bad debts records.



Dashboard 15: The frequency distribution bar charts show bad debts transaction under ER service reported in 2016

## Visual Audit Task 5: Analytical Procedure

 Based on the visual analysis shown in dashboard 14 and 15, we believe that, in the later investigation and testing, auditor should focus more on the ER service occurred in March 2016 and try to find out the reason for the exceptional higher transaction value and bad debts reported in March.

# Questions

- Does service type of the account affect the proportion of transaction value under different transaction type? (Related to Visual Audit Task 4)
- Can we get additional information about the specific hospitals? (Related to Visual Audit Task 1)
- Any data missing? (Mismatch between Account table and Transaction table) Why bad debt transactions have higher mismatch rate? (Related to Visual Audit Task 2)
- Any business rule related to the abnormal hype for the distribution of difference between billing date and transaction date? (3 years, 6 years) (Related to Visual Audit Task 3)

# Future Work

- Improve the framework
- Audit visual evidence and reporting standardization
- Experiments to test the framework