DATASCIENCE PORTFOLIO

Lusy Ambarwati Rochimah

Education

Binus University

Master of Computer Science Concentration: IT Strategic and Infrastructure Management 2014 - 2016

Binus University

Bachelor of Computer Science Concentration: Computer Network 2010 - 2014

Working Experiences

Infomedia (2014 - 2017) Site Engineer

- Troubleshooting, coordinating, maintaining, and monitoring network and system at Telkom Indonesia corporate customer (Bank Mandiri and Grand Indonesia Area).
- Ensuring the implementation of Access Network, Integration & Migration functions to support performance achievement.

Mastersystem Infotama (2017) Customer Relation Manager

- Deals with some or every single issue of the customer's through connecting with them by email, phone, fax, internet or at times interact with them personally
- Works together with the employees and give guidance during complicated situations for them to accomplish the best promising decision for the company and for the clients.
- Make sure costumer adoption of Microsoft Office 365 and Azure Adoption / Consumption
- Maintains professional and technical knowledge by attending educational workshops, reviewing
- professional publications, establishing personal networks, participating in professional societies

LingoAce (2020 - Present) IT Support & Operation Engineer

- Provided technical assistance with computer hardware and software
- Resolved technical issues for clients
 (LingoAce parents or students) via phone,
 in person, or electronically
- Log bugs and enhancement requests
- Performed hardware and software troubleshoot, configurations and updates as needed
- Resolved technical issues in a timely manner using available resources within the company
- Managed and supervise in the usage of company's inventory (e.g., laptop, handphone, headset)
- Setup accounts for Newcomers
- Gathered monthly attendance data from machine
- Be part of a regional team to establish IT operation in the SEA region





Skills and Proficiency

- Fundamentals of Data Science
- Statistics
- Programming knowledge
- Data Manipulation and Analysis
- Data Visualization
- Machine Learning

Data Science Project







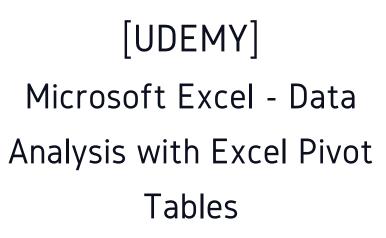
Exploring Machine Learning Models with Python



Incident Management Data Analysis

Certification







[UDEMY]
SQL - MySQL for Data
Analytics and Business
Intelligence



[DIBIMBING]

Data Science Bootcamp



Outline

- Background
- Exploratory Data Analysis & Feature Engineering
- Insight

Incident Management

An "incident" in ITIL (Information Technology Infrastructure Library) terminology defined as:

- an unplanned interruption to IT service
- reduction in quality of IT Service.
- Failure of a configuration item that has not yet impacted.

Incident Management is the process for dealing with all incidents, include:

- failures
- questions or queries reported by the users



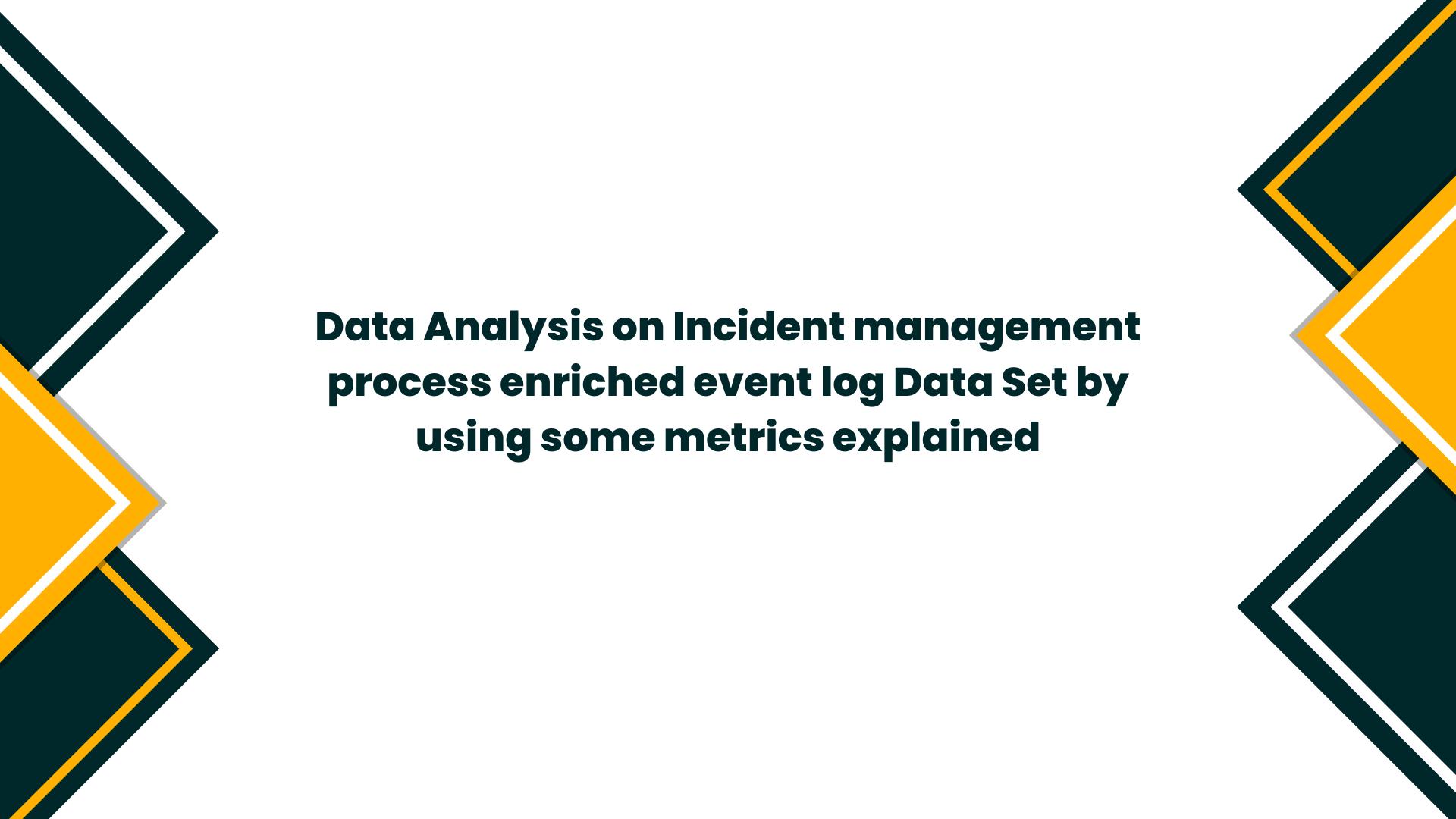
Value to business

- The ability to detect and resolve incidents which results in lower downtime to the business
- The ability to align IT activity to real-time business priorities
- The ability to identify potential improvements to services
- The Service Desk can identify additional service or training requirements found in IT or the business.

Some metrics that should be monitored and reported upon to judge the efficiency and effectiveness

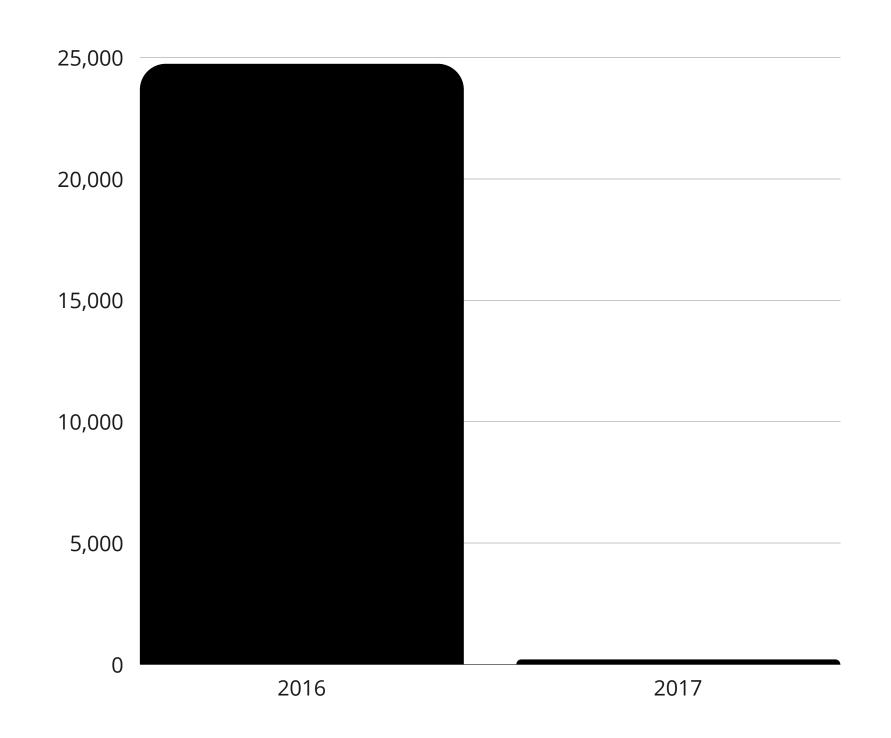
- The Total Number of Problems Recorded in The Period
- Breakdown of incidents at each stage
- Size of daily incident backlog
- Mean elapsed time to achieve incident resolution or circumvention
- Number of incidents reopened and as a percentage of the total
- Number and percentage of incidents incorrectly assigned
- Percentage of incidents handled within agreed response time



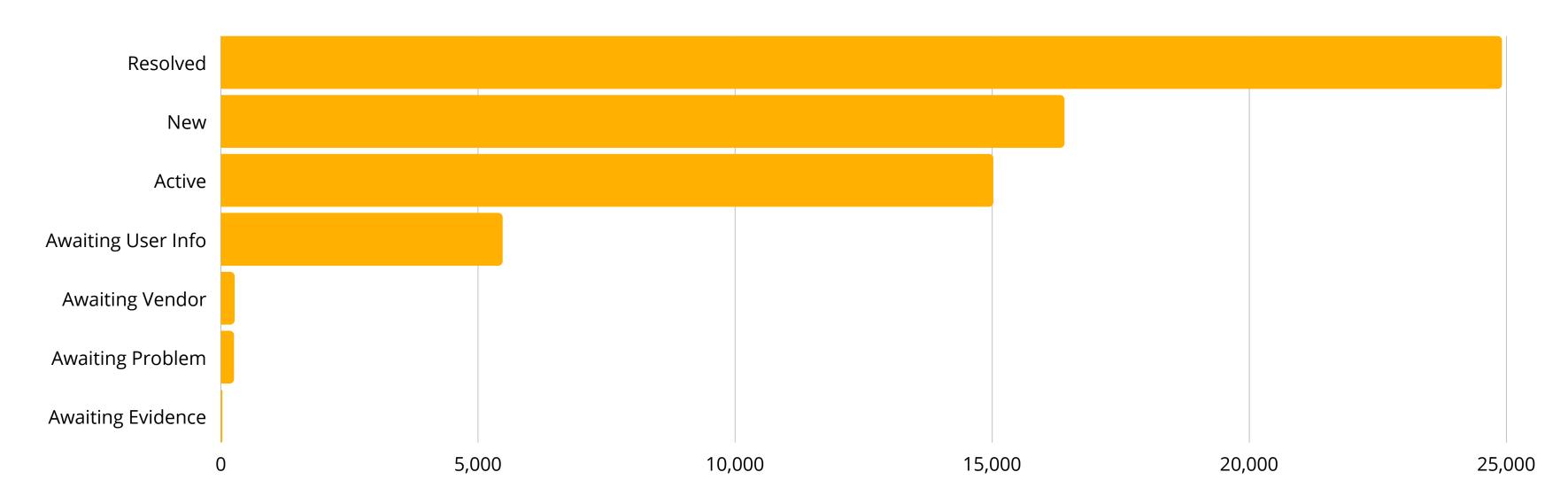


The total number of problems recorded in the period (as a control measure)

Ticket numbers significantly decrease in 2017

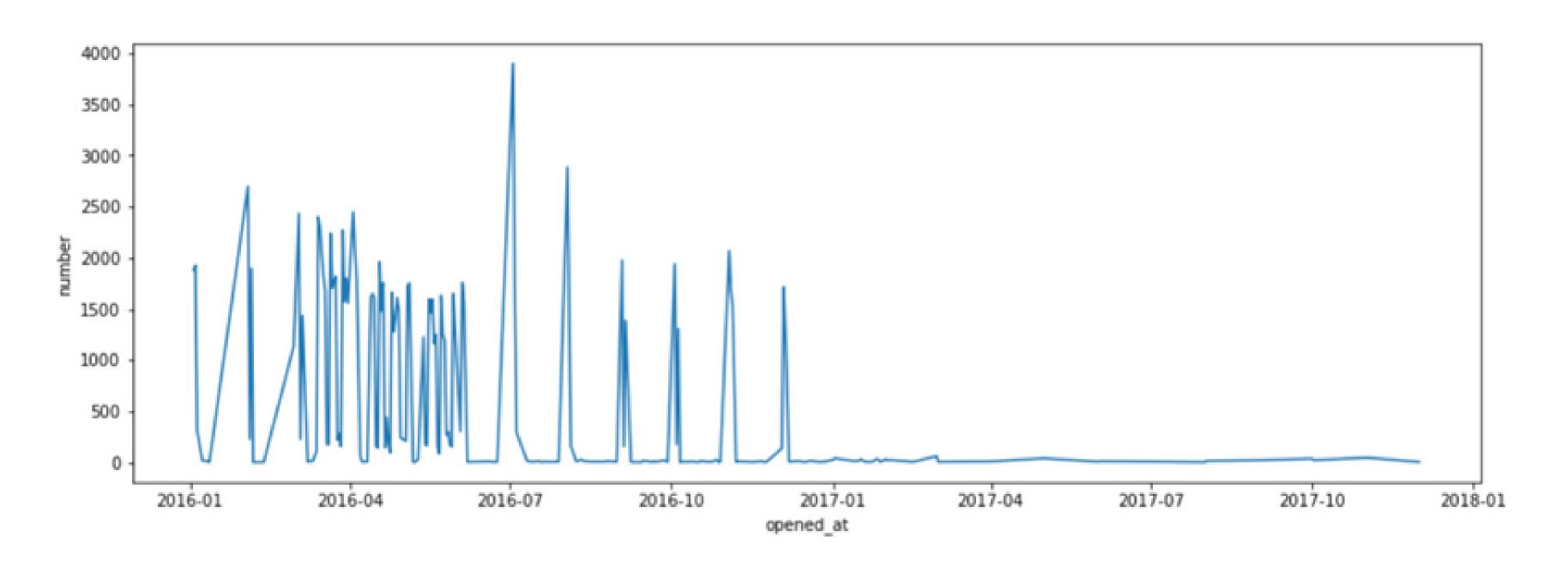


Breakdown of incidents at each stage



From total 24918 ticket number, 24905 tickets resolved

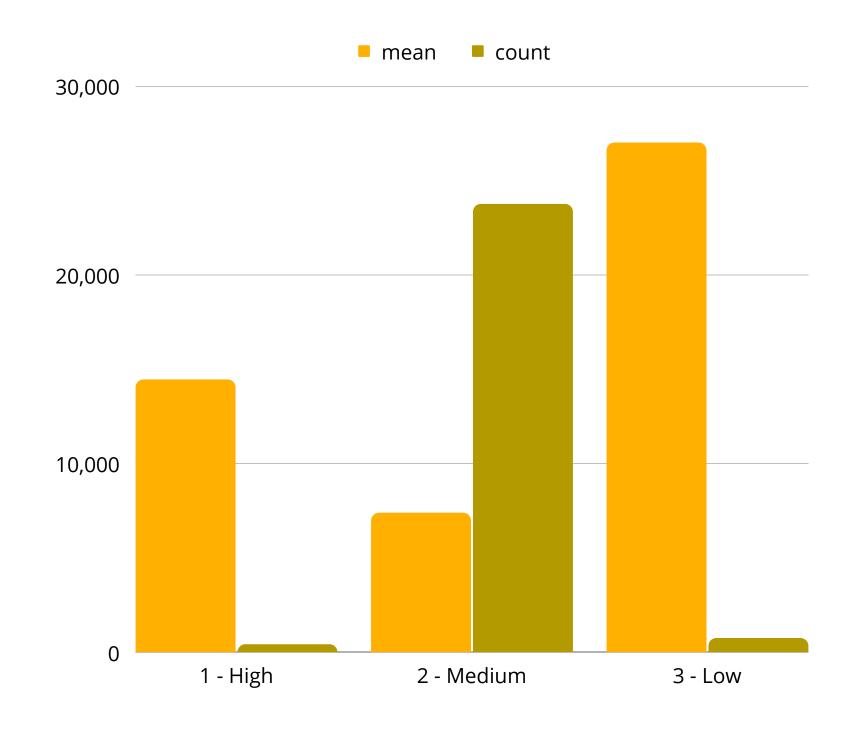
Size of Daily Incident Backlog



The result shown more stable and less ticket backlog in 2017

Mean elapsed time to achieve incident resolution or circumvention, broken down by impact code

'Medium' impact has the lowest mean elapsed time, although it has the most ticket number

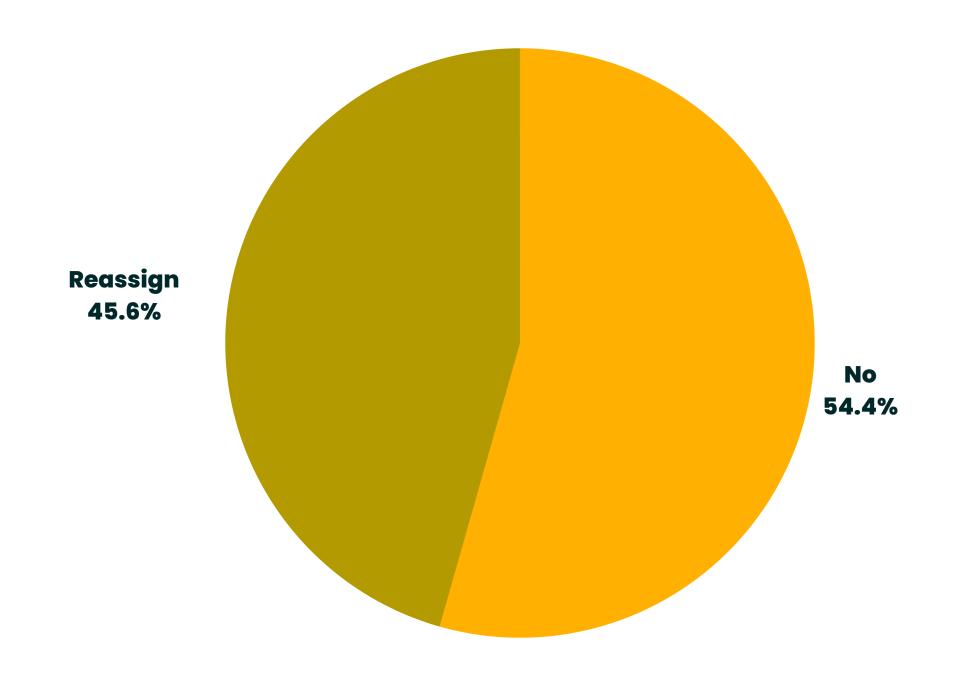


Number of incidents reopened and as a percentage of the total



Only 1.1% ticket reopen

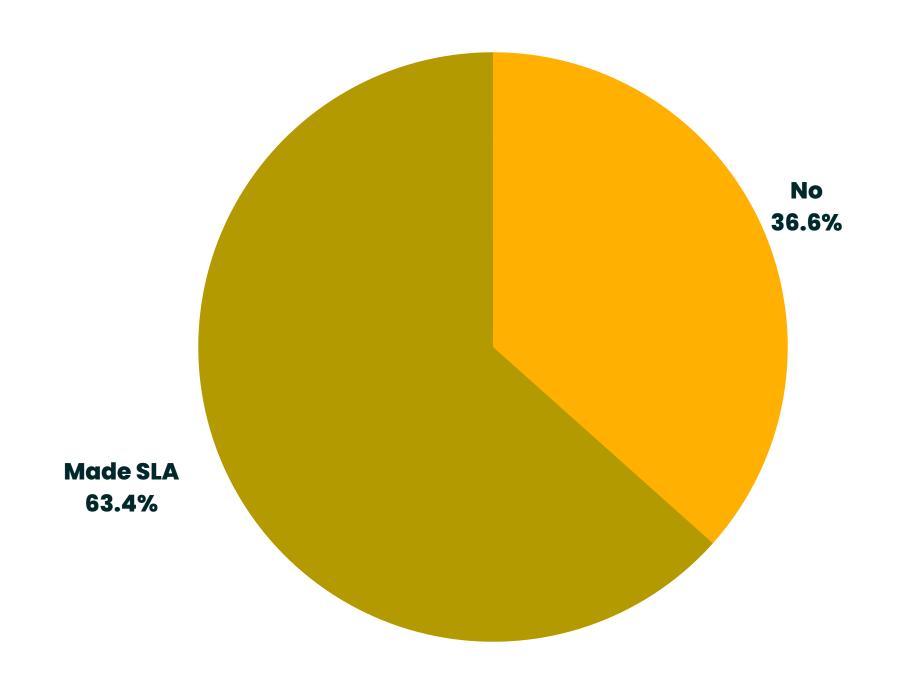
Number and percentage of incidents incorrectly assigned



The result shown 45.6% ticket reassign

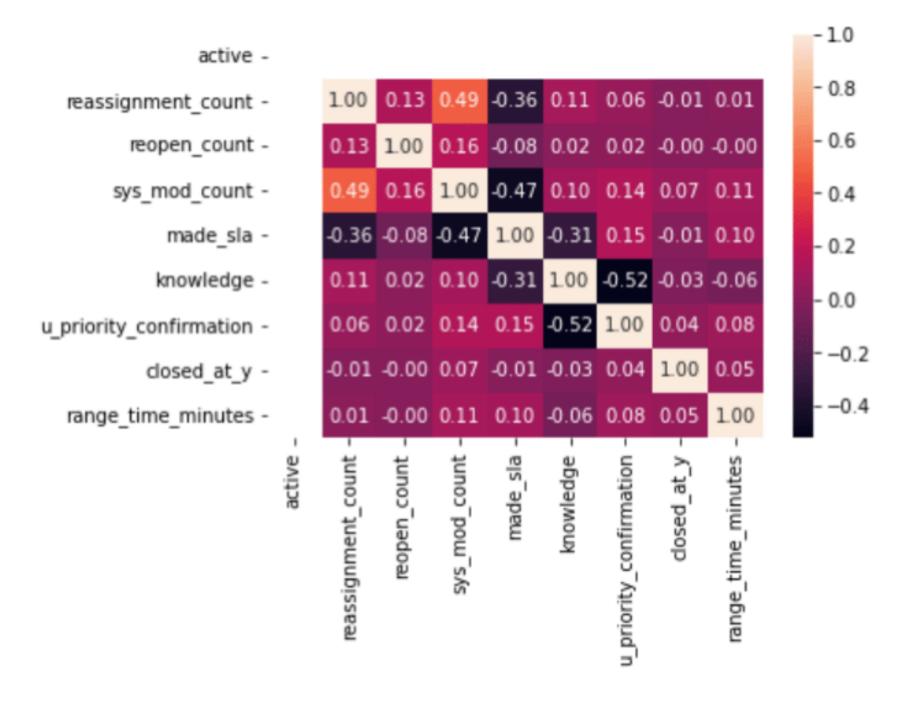
What features are important to predict incidents handled within agreed response time?

Percentage of incidents handled within agreed response time



36.6% ticket not made SLA

Correlation heatmap analysis



Threshold: absolute(corr) >= 0.8
There's no multicollineary



Decision tree comparation between baseline and transformed data

```
print('Accuracy',accuracy_score(y_test, y_pred_dt))
```

Accuracy 0.8966024612092027

Baseline

```
print('Accuracy',accuracy_score(y_test, y_pred_dt))
```

Accuracy 0.9519796682718031

Transformed data

Transformed data has better accuration than baseline

Evaluate other modeling method

Mean absolute error (MAE):

 How far in absolute basis is the model's prediction from the actual data on average

Root Mean Squared Error (RMSE):

 The standard deviation of our prediction errors with respect to the regression line

R2 score:

 In regression, the R2 coefficient of determination is a statistical measure of how well the regression predictions approximate the real data points.

	Method	MAE	RMSE	R2 Score
0	LinearRegression	0.301395	0.396511	0.327398
1	Lasso	0.408038	0.435377	0.189079
2	Ridge	0.274934	0.357140	0.454337
3	ElasticNet	0.389344	0.425065	0.227037
4	RandomForestRegressor	0.062520	0.193162	0.840379
5	XGBRegressor	0.074503	0.152394	0.900646

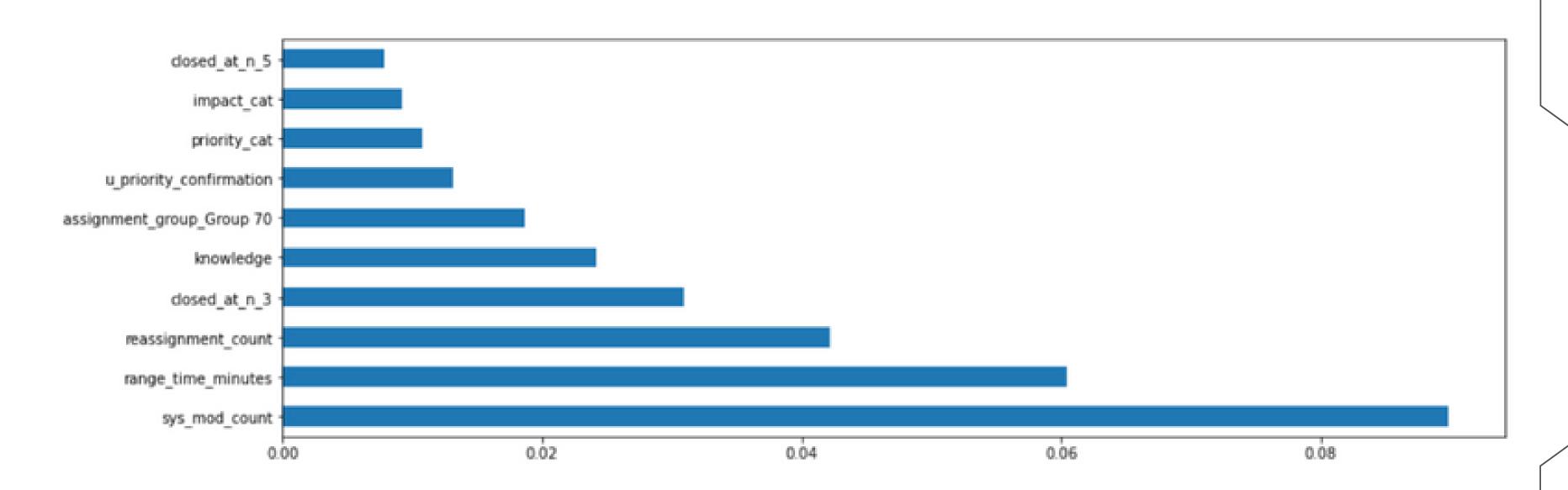
Result shown that RandomForestRegressor and XGBRegressor has better score than other method

Tuning Hyperparameters

Best estimator (bootstrap=False, max_features=10, n_estimators=25)

	params	mean_test_score	rank_test_score
23	{'bootstrap': False, 'max_depth': None, 'max_f	-0.092728	1
21	{'bootstrap': False, 'max_depth': None, 'max_f	-0.098255	2
11	{'bootstrap': True, 'max_depth': None, 'max_fe	-0.098276	3
22	{'bootstrap': False, 'max_depth': None, 'max_f	-0.102785	4
9	{'bootstrap': True, 'max_depth': None, 'max_fe	-0.103766	5
10	{'bootstrap': True, 'max_depth': None, 'max_fe	-0.105507	6
20	{'bootstrap': False, 'max_depth': None, 'max_f	-0.108857	7
8	{'bootstrap': True, 'max_depth': None, 'max_fe	-0.111369	8
18	{'bootstrap': False, 'max_depth': 50, 'max_fea	-0.175573	9
7	{'bootstrap': True, 'max_depth': 50, 'max_feat	-0.175956	10
19	{'bootstrap': False, 'max_depth': 50, 'max_fea	-0.177665	11
6	{'bootstrap': True, 'max_depth': 50, 'max_feat	-0.179695	12
5	{'bootstrap': True, 'max_depth': 50, 'max_feat	-0.195485	13
17	{'bootstrap': False, 'max_depth': 50, 'max_fea	-0.197686	14
16	{'bootstrap': False, 'max_depth': 50, 'max_fea	-0.198015	15
4	{'bootstrap': True, 'max_depth': 50, 'max_feat	-0.199754	16
3	{'bootstrap': True, 'max_depth': 10, 'max_feat	-0.217954	17
2	{'bootstrap': True, 'max_depth': 10, 'max_feat	-0.219017	18
15	{'bootstrap': False, 'max_depth': 10, 'max_fea	-0.219400	19
14	{'bootstrap': False, 'max_depth': 10, 'max_fea	-0.222346	20
12	{'bootstrap': False, 'max_depth': 10, 'max_fea	-0.223452	21
13	{'bootstrap': False, 'max_depth': 10, 'max_fea	-0.224741	22
1	{'bootstrap': True, 'max_depth': 10, 'max_feat	-0.225406	23
0	{'bootstrap': True, 'max_depth': 10, 'max_feat	-0.225921	24

Feature importance



10 features above are the most important feature to predict incidents handled within agreed response time

Insight

- Ticket numbers significantly decrease in 2017 compare than 2016 and the result also shown more stable and less ticket backlog in 2017, it means company does massive improvement that affecting the amount of reporting ticket
- From total 24918 ticket number, 24905 tickets resolved, that's a sign that the company can handle 99,9% well
- 'Medium' impact has the lowest mean elapsed time, although it has the most ticket number, so companies can focus to find out what causes 'High' and 'Low' impact has higher elapsed time

Insight

- Only 1.1% ticket reopen, it is shown that the ticketing agent understand enough when they can close the ticket
- The result shown 45.6% ticket reassign, it could impacting mean elapsed time, companies have to find out what causes this problem (e.g. SOP that not clear enough, agent understanding related this case)
- 36.6% ticket not made SLA, it's a huge amount, 10 most important feature predict incidents handled within agreed response time can help companies to improve ticket that made SLA



Source

Data Set: Incident management process enriched event log Data Set by UCI Machine Learning Library. https://archive.ics.uci.edu/ml/datasets/Incident+management+pr

ocess+enriched+event+log#

Reference

Commerce, Office of Government. (2007). The Official Introduction to the ITIL Service Lifecycle. United Kingdom: Office of Government Commerce.

