

# Applying Process Mining to a Dispatching Process

Presenter Name: Abs Amiri

Date: 10-June-2016

- *Professionalism*
- *Teamwork*
- *Innovation*

# Disclaimer

---

- ❑ I am here representing myself and the projects that I have initiated and worked on.
- ❑ By no means I am representing these companies and therefore no liabilities expected on them.

# Outline

---

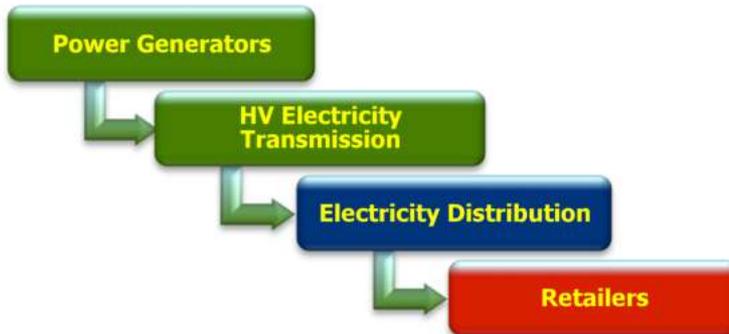
- ❑ About me.
  
- ❑ About SPARKLE
  
- ❑ Case Study (Dispatching Process)
  - In brief about Dispatching Process.
  - How and where the idea initiated.
  - Efficiency Targets.
  - Outcome
  
- ❑ Process Mining penetration strategy & plan in the organisation.
  
  
- ❑ Summary
- ❑ Questions?

# Introduction

---

## □ SPARKLE

- ICT services to Queensland's electricity suppliers



# Dispatching Process



# Dispatching Process



# The Drives and Needs for **PROCESS IMPROVEMENTS**

---

**Government & Public pressure (Customer centricity)**

**The increased interest in the area of process improvement**

**Competition. (efficiency is often one of the requirements for competitiveness)**

**The effectiveness of dispatching emergency orders during special weather events**

**Avoid bottlenecks, idle times, and delays in business processes**

**Quicker power restoration (STIPIS reward)**

"The identification of these problems, which may affect the efficiency of a process, can be carried out through a process mining exercise. Similarly to other organisations, we are interested in knowing if our processes are efficient and effective"

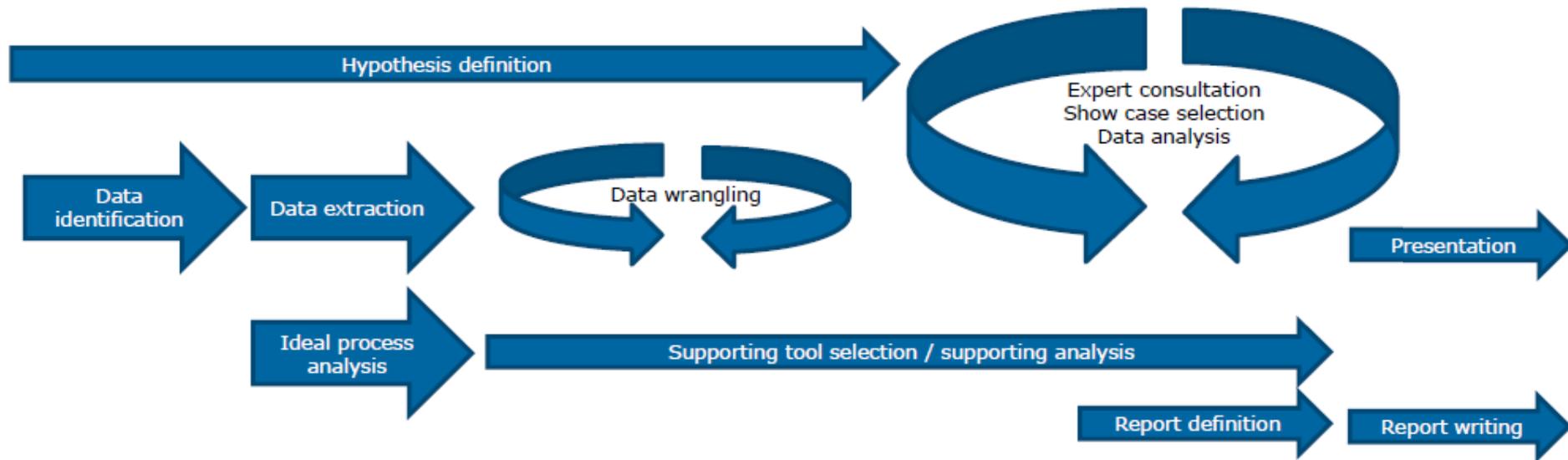
# Process Mining Scope

---

- ❑ 4 Data Sources for Dispatching Process
  - Call Centre.
  - DMS (PON).
  - Ventyx Service Suite 9.2.
  - Dispatch groups phone logs.
  
- ❑ Only complete processes (end-to-end) captured for the analysis.
  
- ❑ Sample of data (Emergency Orders) extracted/prepared for the analysis (10-11 Dec 2015) Special weather event.

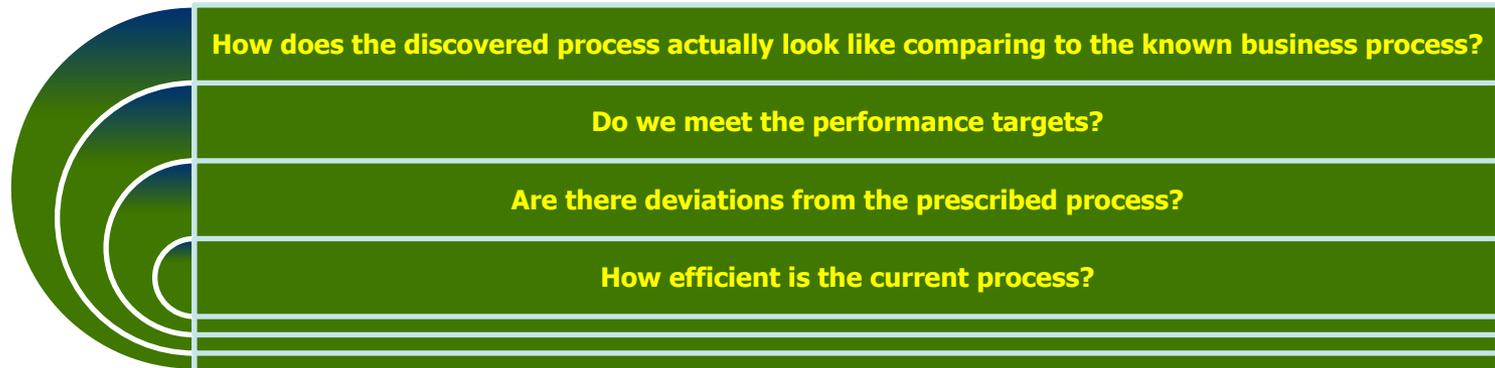
# Project Approach

---



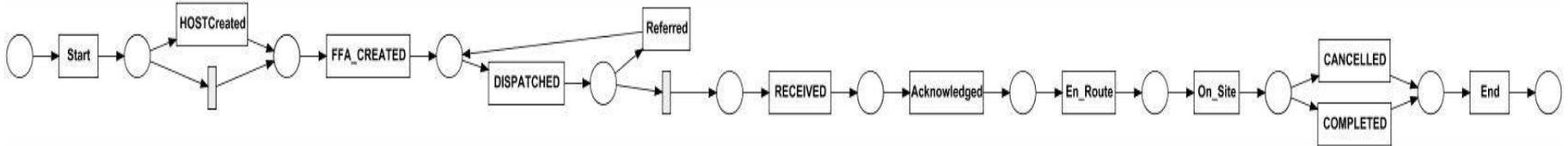
# Process Mining Questions

---

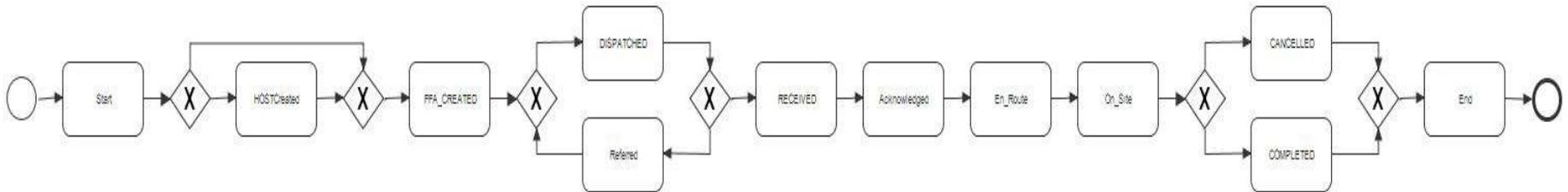


# Process Discovery

- As is process discovered (Petri-Net model/Alpha-Algorithm)



- As is process discovered (Process Tree as BPMN model)



# Data Challenges

---

- ❑ Tooling & technicalities
  - ETL C# is used (not a typical ETL tool).
  - Oracle Database/SQL Scripts.
- ❑ Data issues
  - **Timestamps are implicitly available**

# Timestamps Implicitly Available

## ❏ From-to records

- Database records that status changes from A to B at time T
- For PROM, DISCO, R Studio & RapidMiner:
  - Select time in and time out per activity
  - Deduce additional timestamps for start and end of process

Database

Case	Timestamp	Change
10001	12-6-2015 9:22	New case
10001	12-6-2015 10:34	Status change to B
10001	12-6-2015 11:46	Status change from B to C
10001	12-6-2015 12:58	Status change from C to D
10001	12-6-2015 14:10	Status change from D to E
10001	12-6-2015 15:22	Case closed

Disco

Case	Activity	Time In	Time out
1001	A	12-6-2015 9:22	12-6-2015 10:34
1001	B	12-6-2015 10:34	12-6-2015 11:46
1001	C	12-6-2015 11:46	12-6-2015 12:58
1001	D	12-6-2015 12:58	12-6-2015 14:10
1001	E	12-6-2015 14:10	12-6-2015 15:22

# Data Challenges

---

- ❑ Tooling & technicalities
  - ETL C# is used (not a typical ETL tool).
  - Oracle Database/SQL Scripts.
- ❑ Data issues
  - Timestamps are implicitly available
  - **Redundant timestamps**

# Redundant Timestamps

---

## □ Grouped timestamps

- If timestamps of events are very close to each other, and the records together describe in fact 1 “real life event”, the Algorithm still assigns meaning to the order of these events.
- Easiest way to deal with this is to filter out all non-essential activities in PROM & Disco.
- Alternative is to create an omnibus activity with the minimum time in and maximum time out of relevant activities (before loading the data).

Case	Activity	Time In	Time out
1001	Close phase	12-6-2015 9:22	12-6-2015 10:34

# Data Challenges

---

## ❑ Tooling & technicalities

- ETL C# is used (not a typical ETL tool).
- Oracle Database/SQL Scripts.

## ❑ Data issues

- Timestamps are implicitly available
- Redundant timestamps
- **Combination of activities and statuses**

# Combination of Activities and Statuses

---

## □ Activities and statuses

### ○ Data wrangling challenge

- Activities only have time in, Statuses have time and time out.
- Activities happen during Statuses.

### ○ PROM & DISCO:

- Mixing Status and Activity information does not give expected results.
- For now, we often use start times only

# Data Challenges

---

## ❑ Tooling & technicalities

- ETL C# is used (not a typical ETL tool).
- Oracle Database/SQL Scripts.

## ❑ Data issues

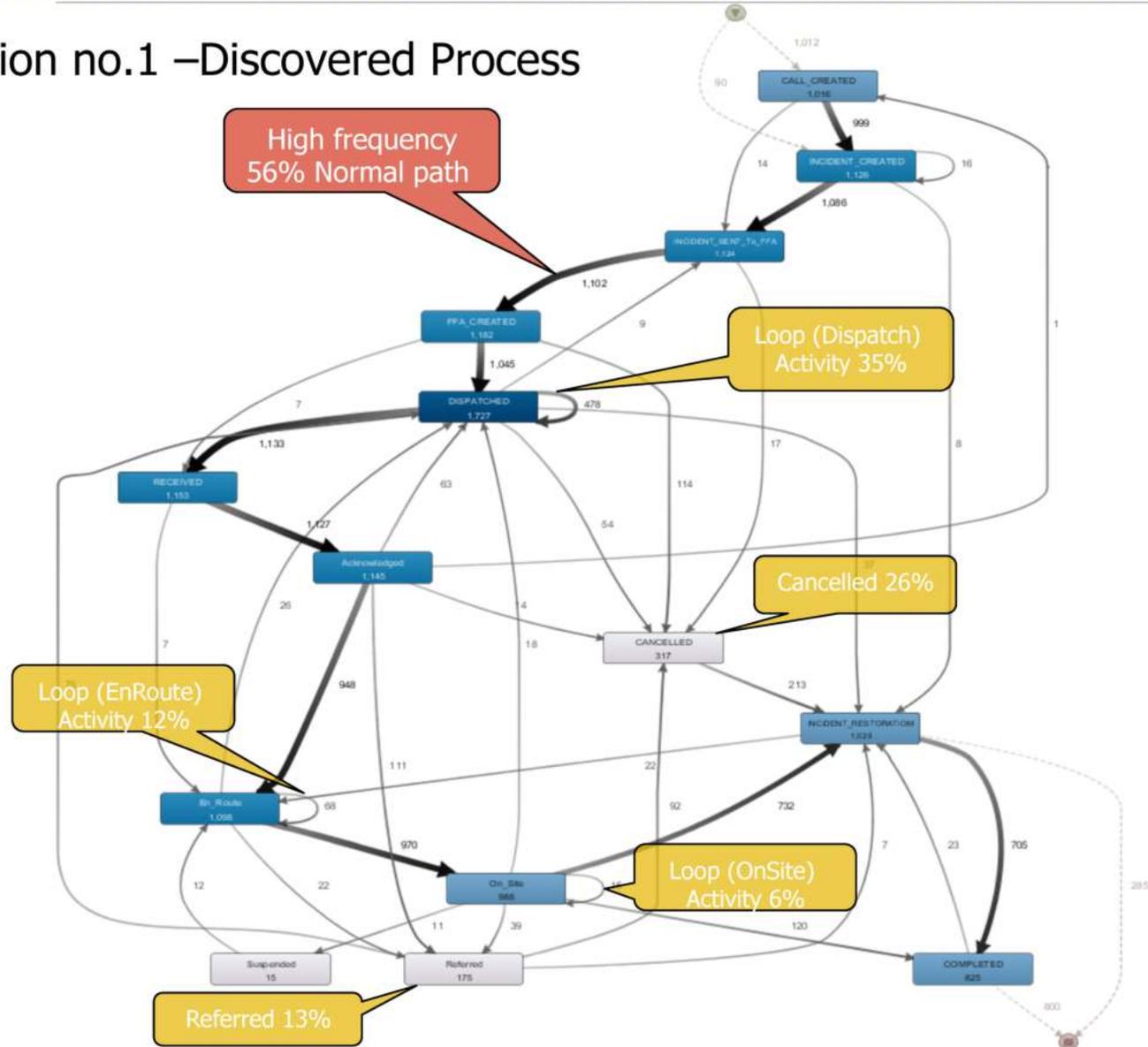
- Timestamps are implicitly available
- Redundant timestamps
- Combination of activities and statuses
- **Information in free text fields/XML**
- **Multiple database tables with varying data formats (integer/text)**
- **Multiple database tables with non-matching content**
- **Timestamps in various time zones and various formats**

## ❑ Format for advanced analysis

- **-Split/add columns**

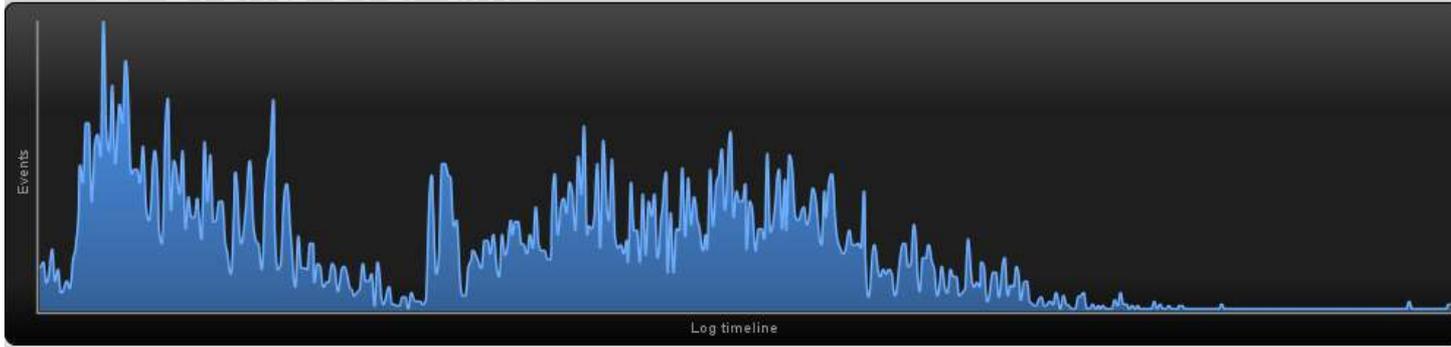
# Overall Process Map

## Question no.1 –Discovered Process



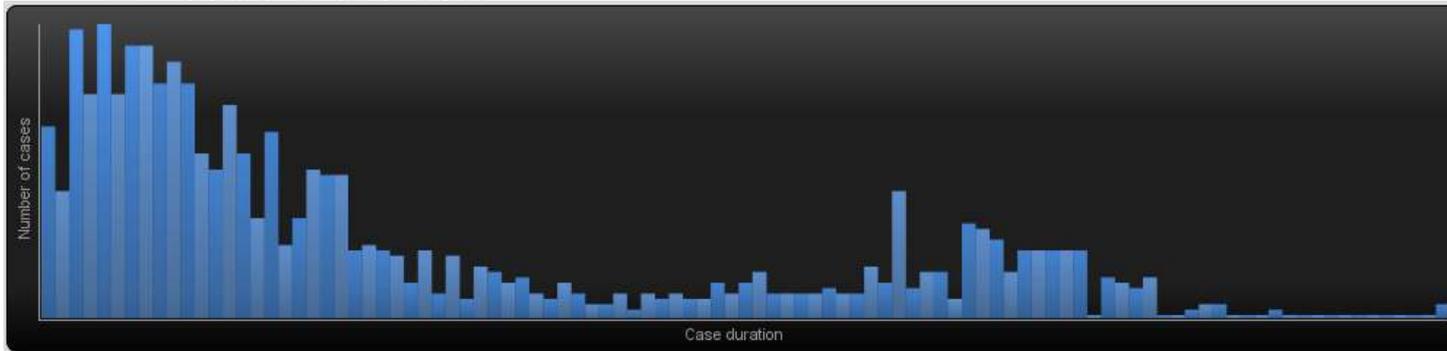
# Overall View

## Events over time



Events	12,920
Cases	1,188
Activities	14
Median case duration	4.4 hrs
Mean case duration	7.6 hrs
Start	10.12.2015 16:00:15
End	12.12.2015 14:01:55

## Case Duration



Events	12,920
Cases	1,188
Activities	14
Median case duration	4.4 hrs
Mean case duration	7.6 hrs
Start	10.12.2015 16:00:15
End	12.12.2015 14:01:55

# Variants grouped by number of cases

## Question No. 2 – Performance Targets

**Threshold**  
 12% <= 1 hr  
 36% <= 3 hrs  
 55% <= 5 hrs  
 31% >= 8 hrs

Variant	Cases	Events	Median duration	Mean duration
Variant 1	378	11	2 hours, 31 mins	3 hours, 41 mins
Variant 2	127	12	5 hours, 53 mins	9 hours, 14 mins
Variant 3	83	6	18 hours, 38 mins	15 hours, 15 mins
Variant 4	41	10	6 hours, 18 mins	10 hours, 55 mins
Variant 5	32	7	2 hours, 17 mins	3 hours, 4 mins
Variant 6	24	13	9 hours, 25 mins	11 hours, 45 mins
Variant 7	19	14	3 hours, 42 mins	4 hours, 20 mins
Variant 8	18	4	34 mins, 2 secs	1 hour, 16 mins
Variant 9	14	7	15 hours, 59 mins	15 hours, 49 mins
Variant 10	11	8	5 hours, 19 mins	9 hours, 14 mins
Variant 11	10	13	4 hours, 54 mins	7 hours, 34 mins
Variant 12	10	12	3 hours, 20 mins	4 hours, 12 mins
Variant 13	9	3	2 hours, 35 mins	4 hours, 6 mins
Variant 14	8	11	6 hours, 2 mins	6 hours, 30 mins
Variant 15	8	13	3 hours, 46 mins	5 hours, 48 mins
Variant 16	8	15	6 hours, 59 mins	7 hours, 7 mins
Variant 17	6	11	4 hours, 21 mins	4 hours, 48 mins
Variant 18	7	15	10 hours, 45 mins	12 hours, 15 mins
Variant 19	7	7	2 hours, 58 mins	4 hours, 7 mins
Variant 20	7	12	3 hours, 51 mins	5 hours, 18 mins

# General Stats..

- General Stats.
- Activities

Activity	▲ Frequency	Relative frequency
DISPATCHED	1,727	13.37 %
FFA_CREATED	1,182	9.15 %
RECEIVED	1,153	8.92 %
Acknowledged	1,145	8.86 %
INCIDENT_CREATED	1,126	8.72 %
INCIDENT_SENT_To_FFA	1,124	8.7 %
En_Route	1,098	8.5 %
INCIDENT_RESTORIOM	1,029	7.96 %
CALL_CREATED	1,016	7.86 %
On_Site	988	7.65 %
COMPLETED	825	6.39 %
CANCELLED	317	2.45 %
Referred	175	1.35 %
Suspended	15	0.12 %

- Top Dispatched areas

Value	▲ Frequency	Relative frequency
MSBLH	4,081	31.59 %
SCBDT	3,293	25.49 %
MSCVL	2,006	15.53 %
WEIPW	382	2.96 %
WEBNH	320	2.48 %
MSGSP	241	1.87 %

- Top Suburbs

Value	▲ Frequency	Relative frequency
JIMBOOMBA	1,854	14.35 %
LOGAN VILLAGE	1,019	7.89 %
MACLEAY ISLAND	781	6.04 %
CEDAR VALE	414	3.2 %
RUSSELL ISLAND	340	2.63 %
REDLAND BAY	292	2.26 %
TAMBORINE	279	2.16 %
GREENBANK	249	1.93 %

# General Stats..

## Hosts

Value	▲ Frequency	Relative frequency
EMERGENCY	12,621	97.69 %
DUPLICATE	299	2.31 %

## Priorities

Value	▲ Frequency	Relative frequency
1_Emergency - Critical	5,169	40.01 %
2_Emergency - Non-critical	5,123	39.65 %
3_Emergency - Lowest	2,628	20.34 %

## Job Codes

Value	▲ Frequency	Relative frequency
D-WDPH_Wires Down Pole to House	2,427	18.78 %
D-WDPP_Wires Down Pole to Pole	2,084	16.13 %
D-LOS_Loss of Supply	1,656	12.82 %
D-MAIN_Tree in Mains	776	6.01 %
D-LO_L/O - 11kv Lockout	630	4.88 %
D-SERV_Tree in Service	585	4.53 %
D-REEN_Reenergisation(after Disconnect)	514	3.98 %
D-RPSY_Repair Job Minor Involving LOS	392	3.03 %
D-URPH_Urgent Public Hazard	377	2.92 %

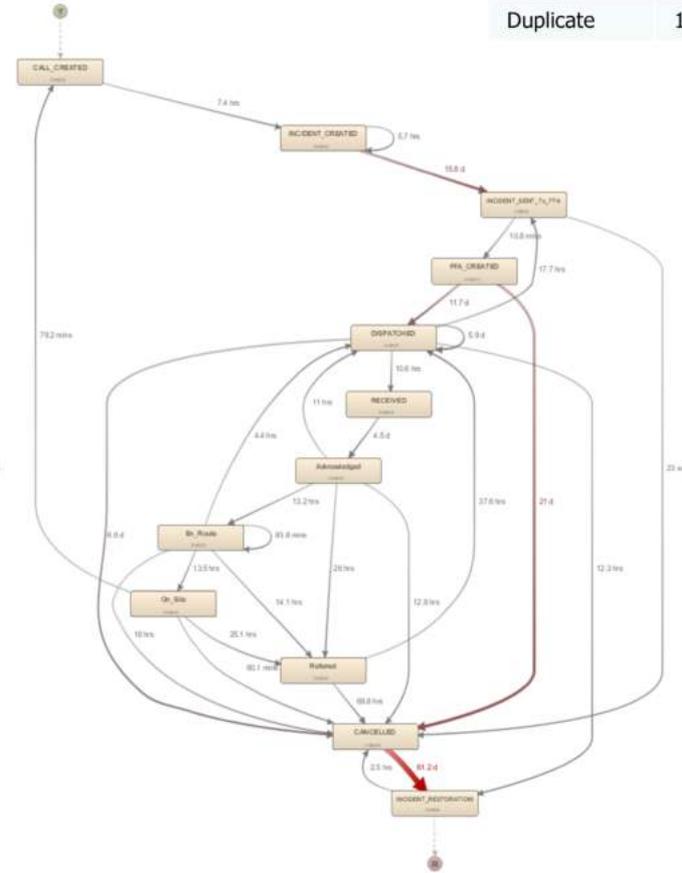
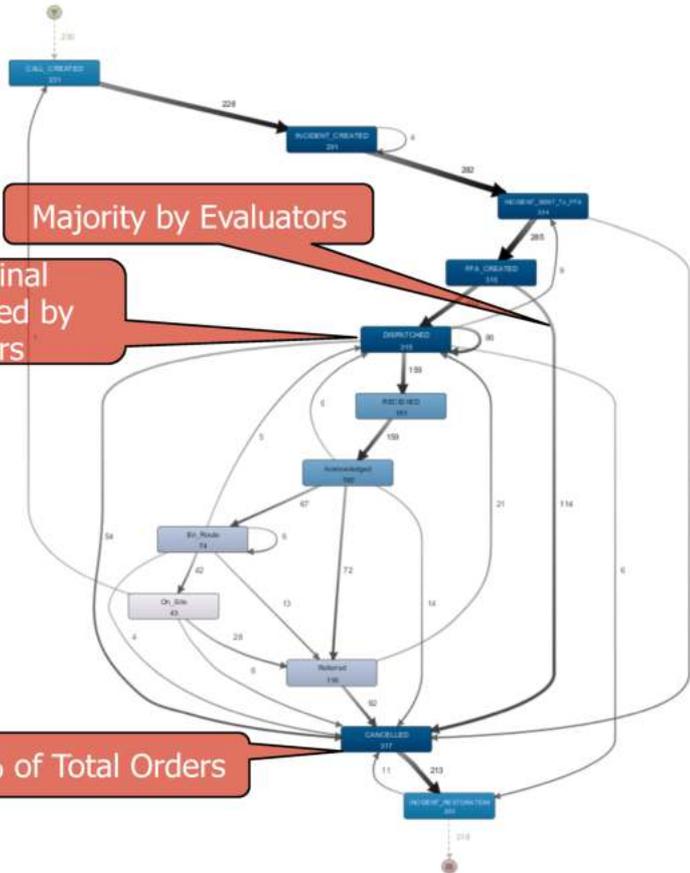
## Business units

Value	▲ Frequency	Relative frequency
011_Field Services	8,865	68.61 %
013_Network Operations	3,320	25.7 %
051_Formway South	419	3.24 %
001_Customer Connections	258	2 %
002_Customer Services	58	0.45 %

# Cancelled orders

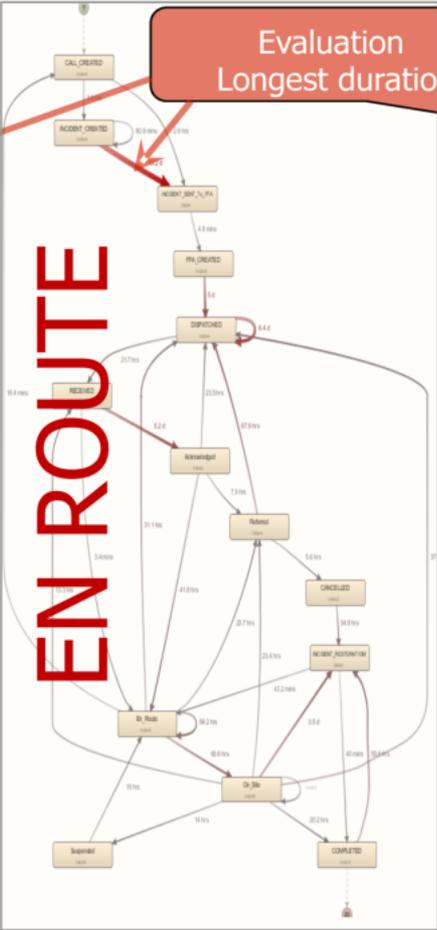
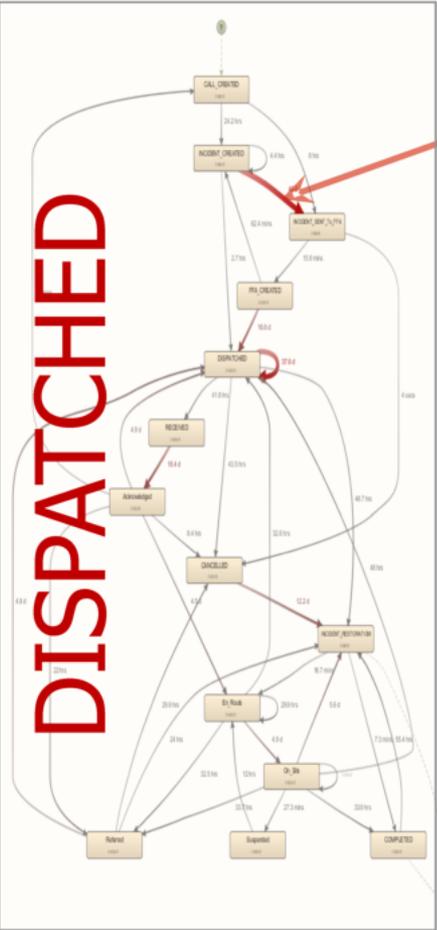
## Question No. 3 – Process Deviations

Value	Freq.	Relative Freq.
Emergency	2453	95.93%
Duplicate	104	4.07%

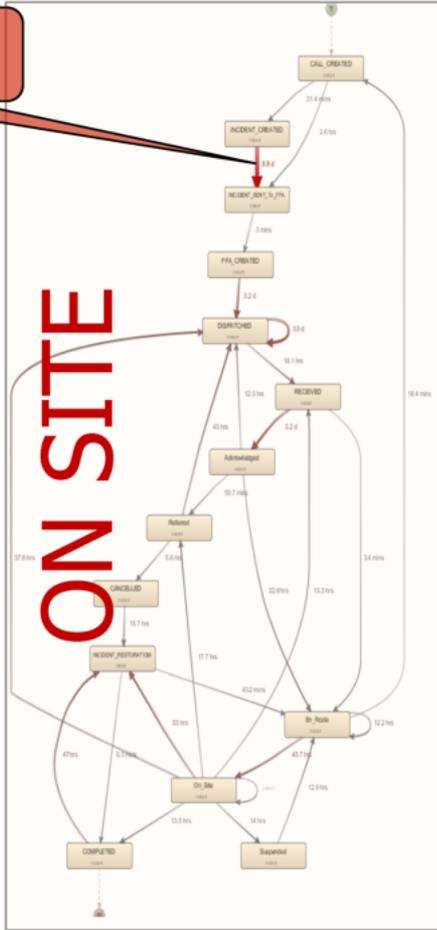


# Activity Loops

Question No. 4 – Process efficiency



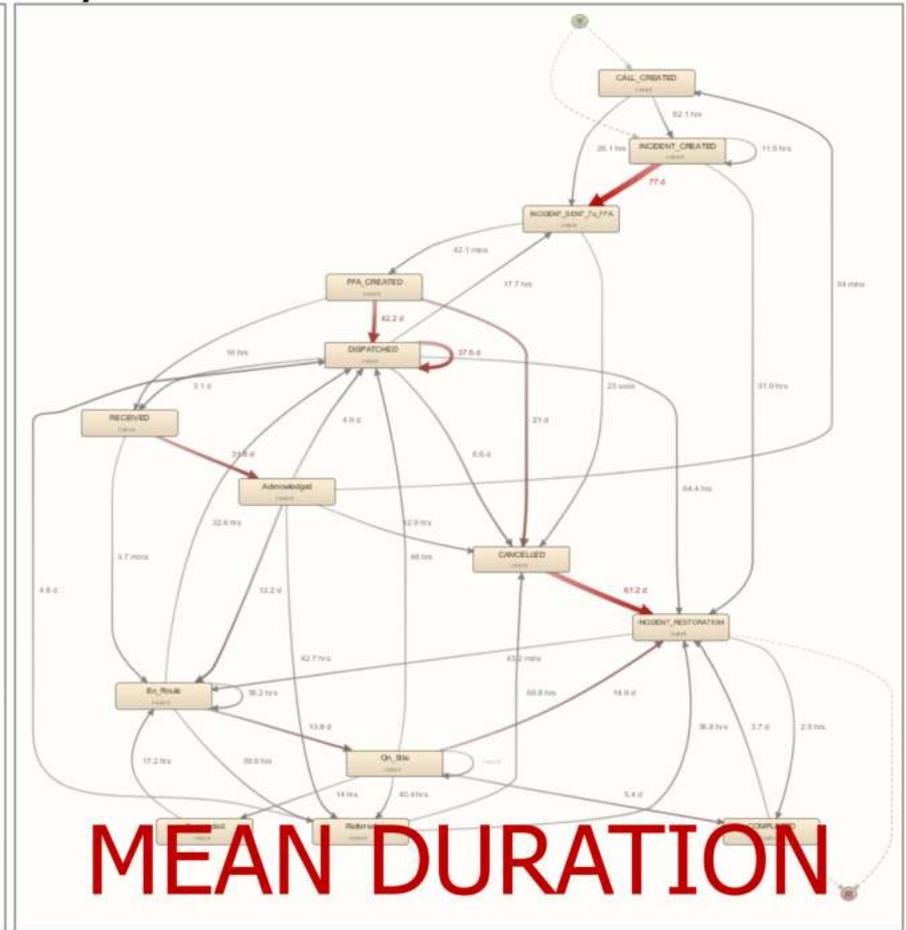
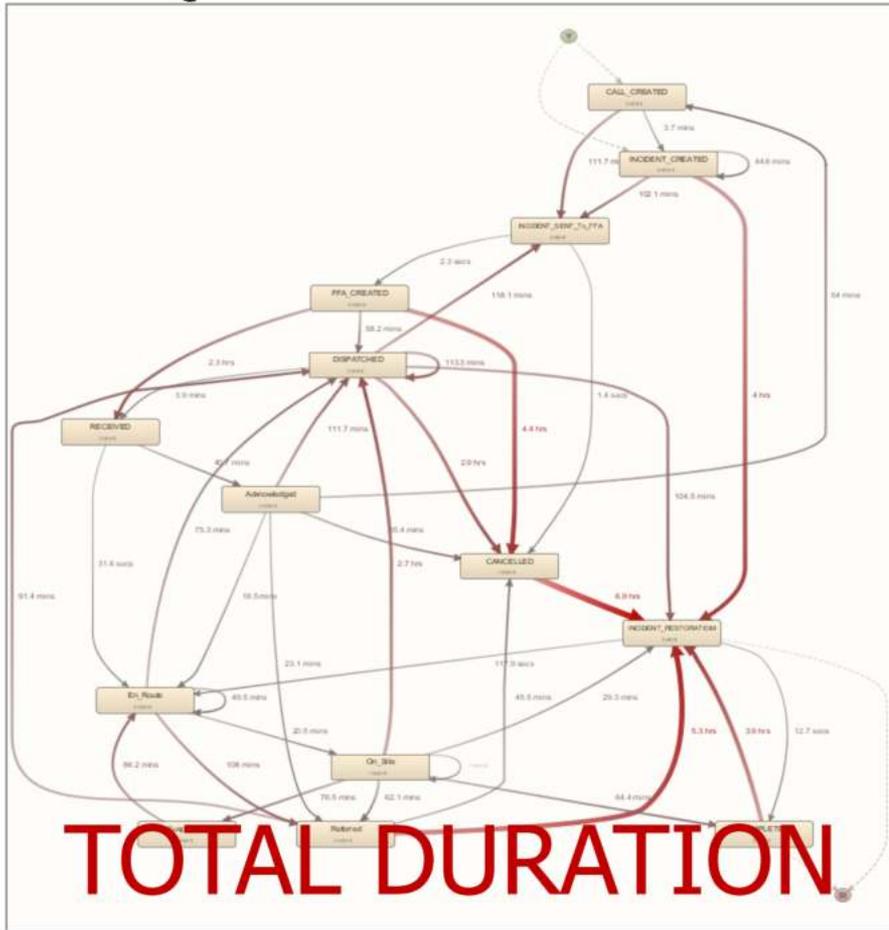
Evaluation  
Longest duration



Activity	Relative Freq.
Dispatched	35%
En Route	12%
On Site	6%

# Performance Analysis

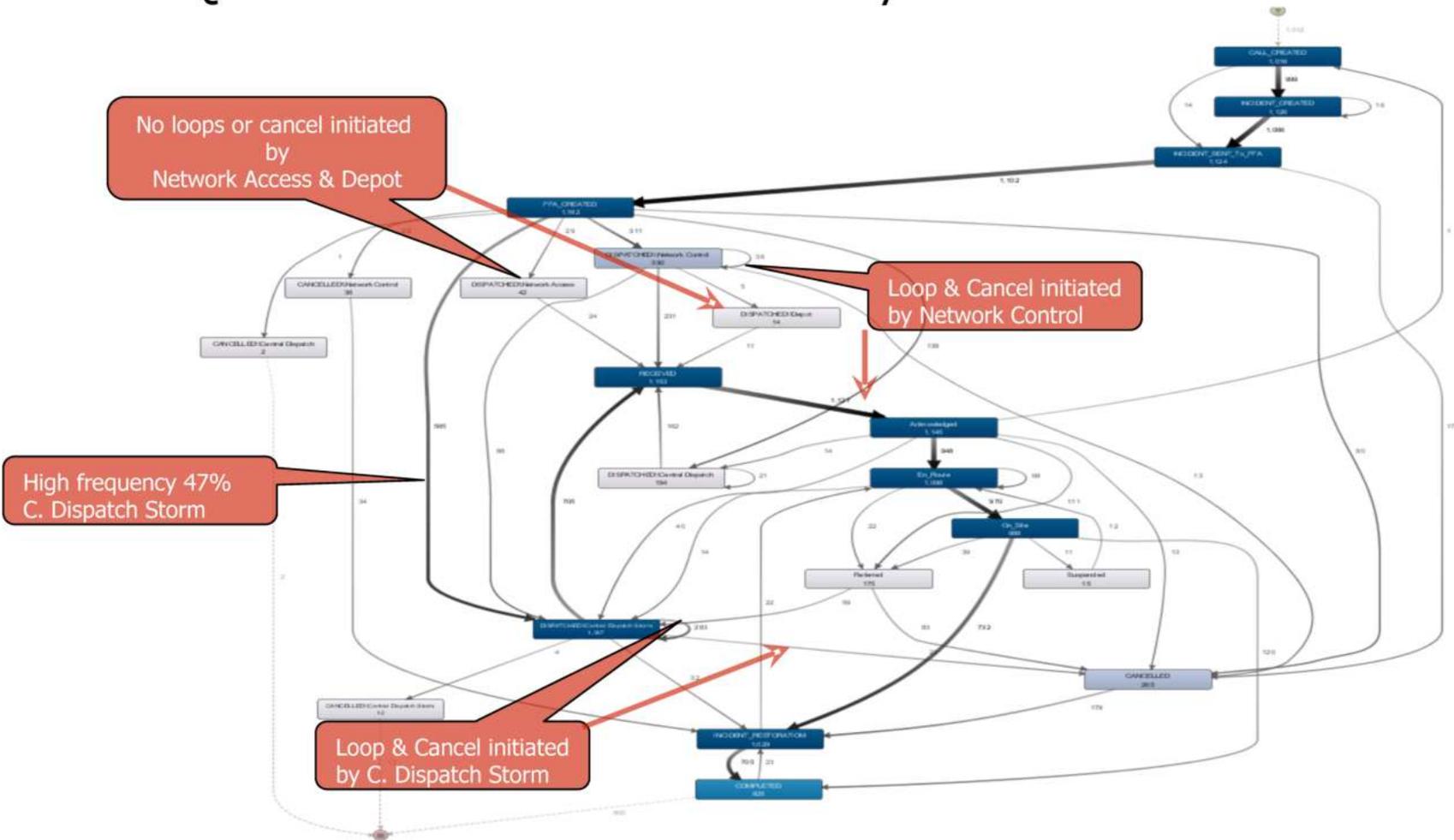
## Question No. 4 – Process efficiency





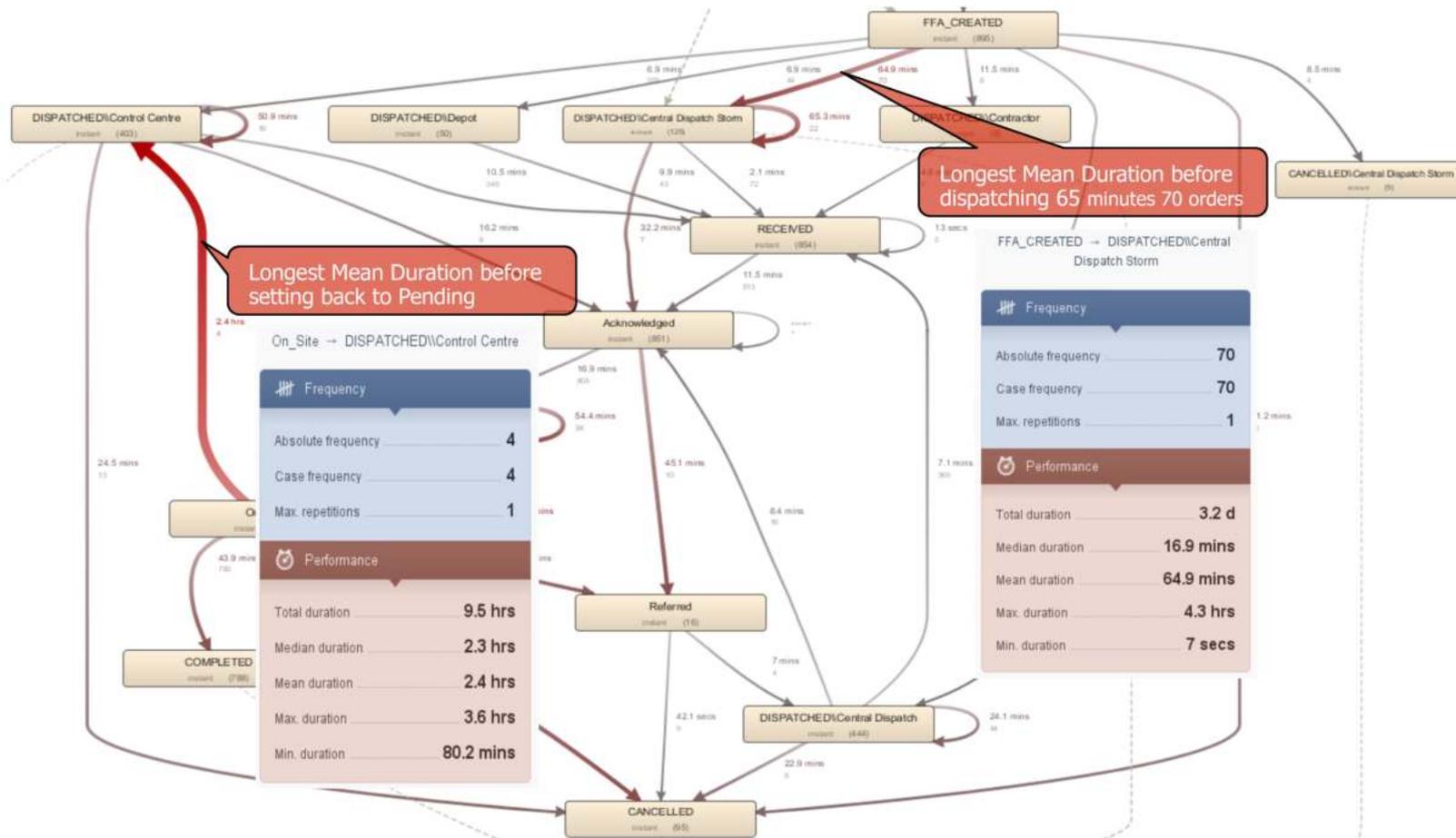
# Dispatched by different groups

## Question No. 4 – Process efficiency

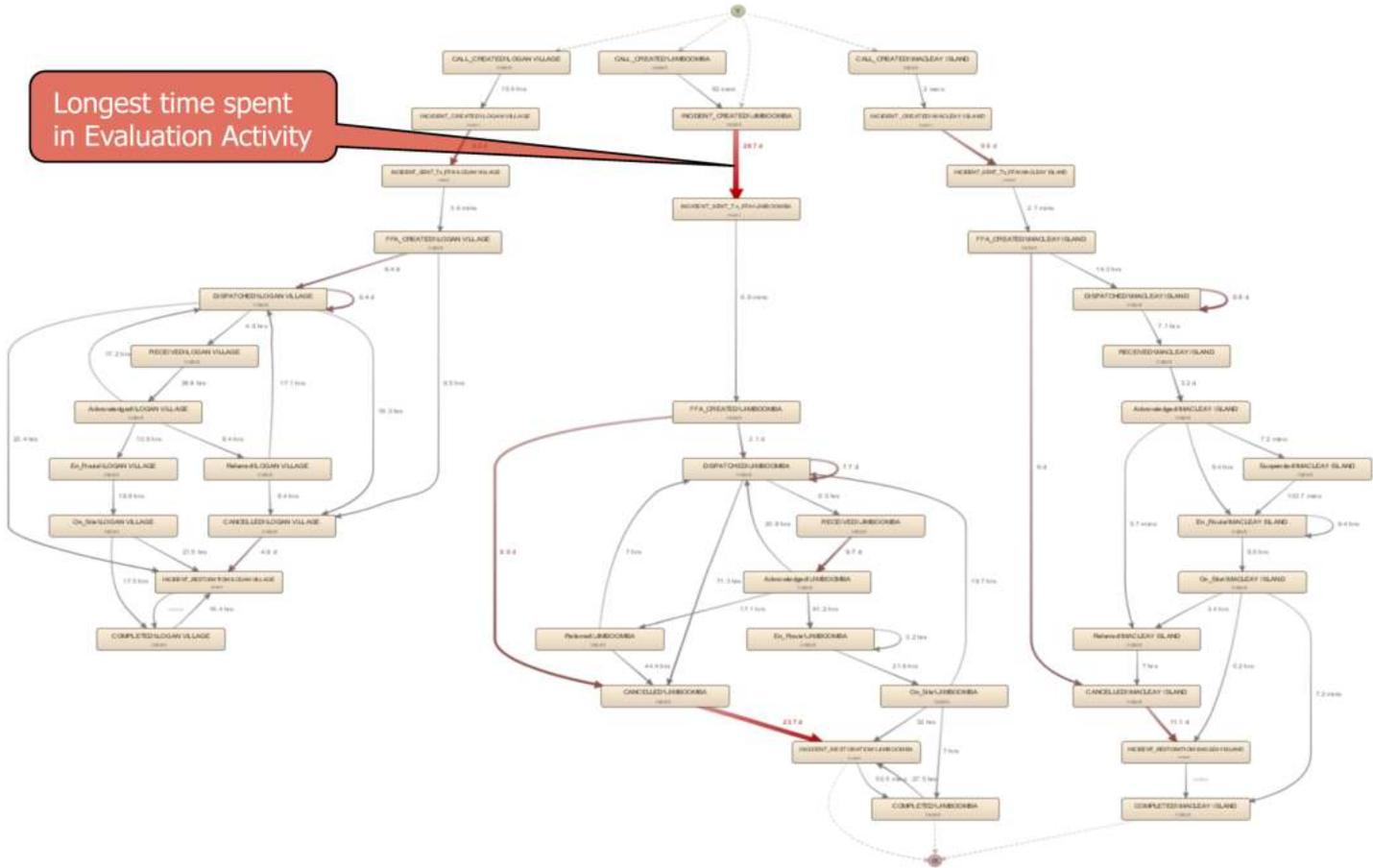


# Dispatched by different groups

## Question No. 4 – Process efficiency

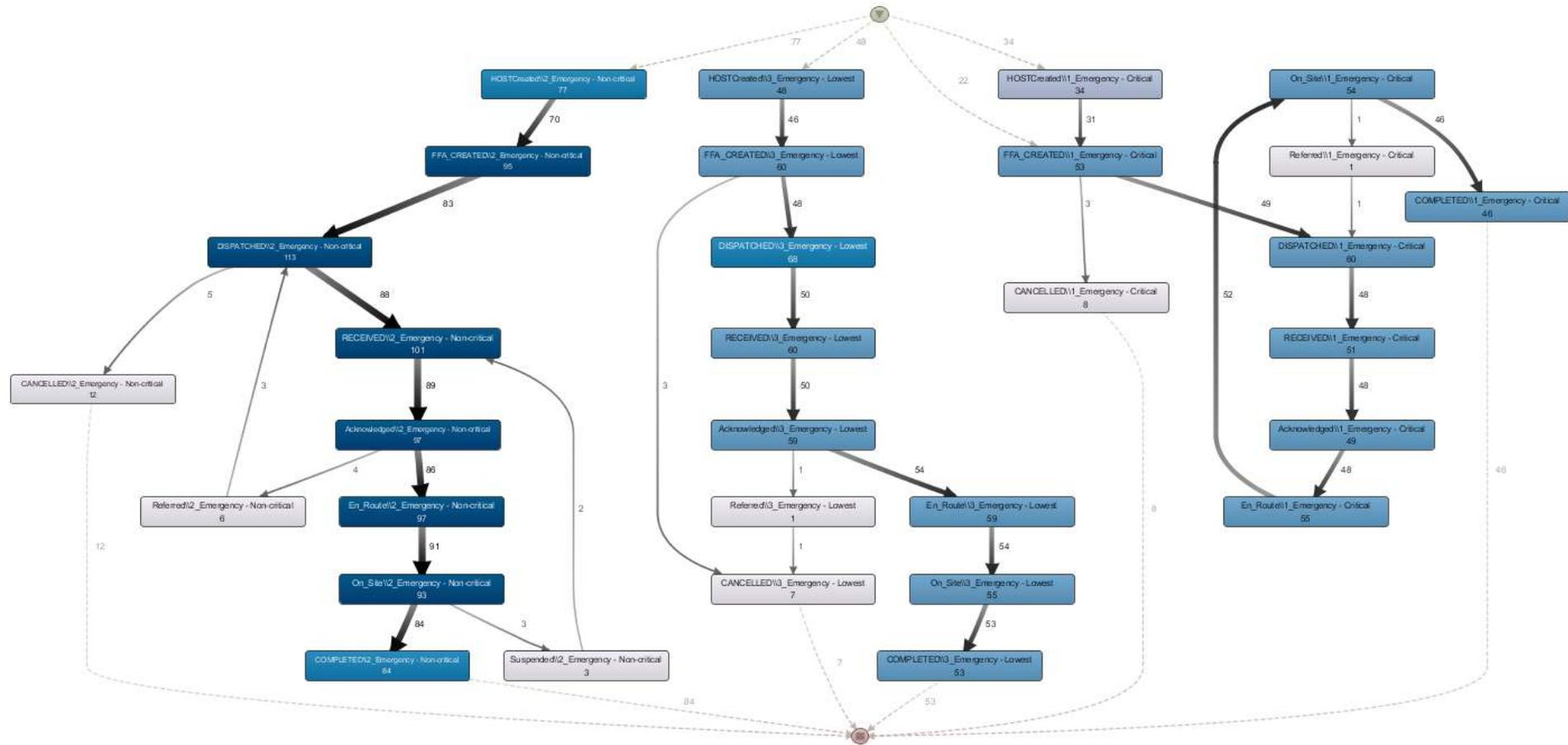


# Top 3 impacted Suburbs



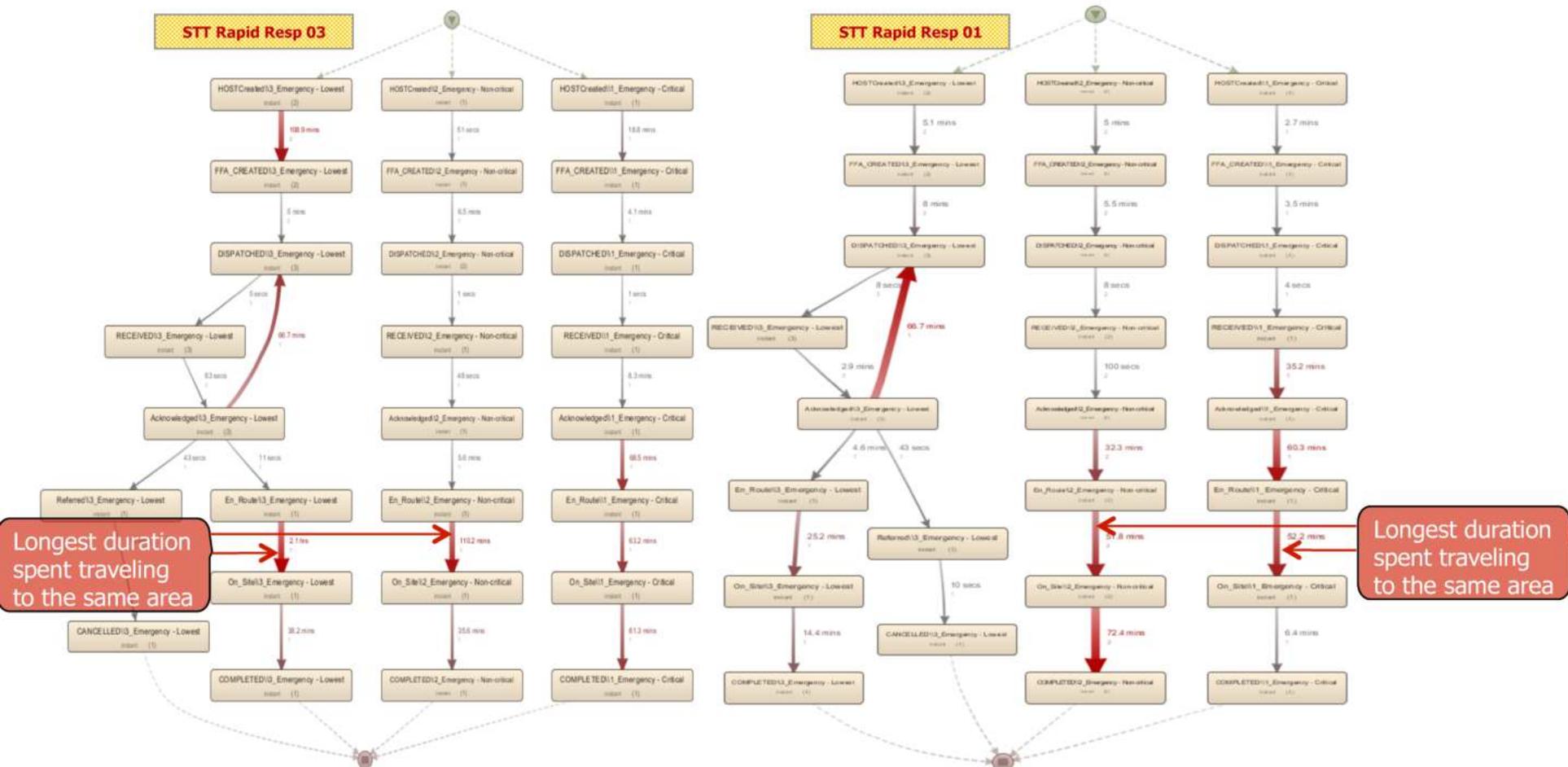
# Process/Data Mining Analysis continue..

## Orders Grouped by Priority (Top 2 areas [MSBLH] [MSGSP] 18% of total orders)



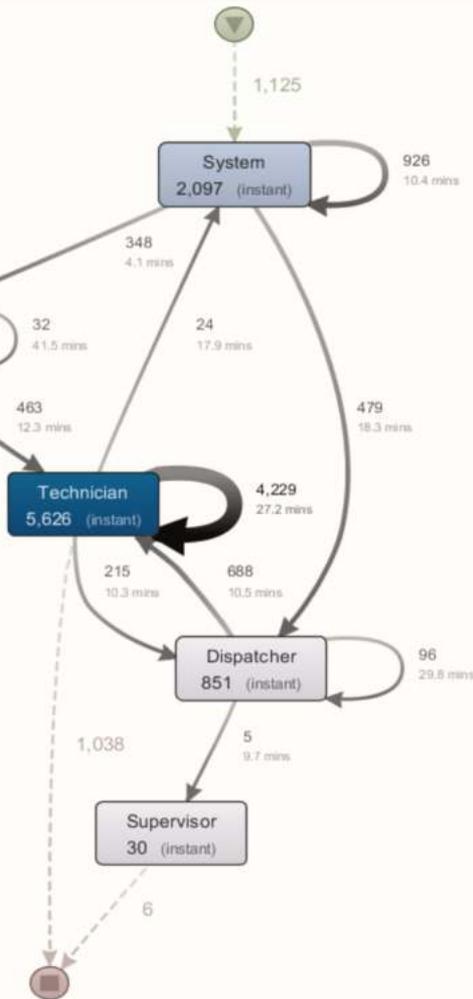
# Orders Grouped by Priority

Areas [MSBLH] (during business hours)

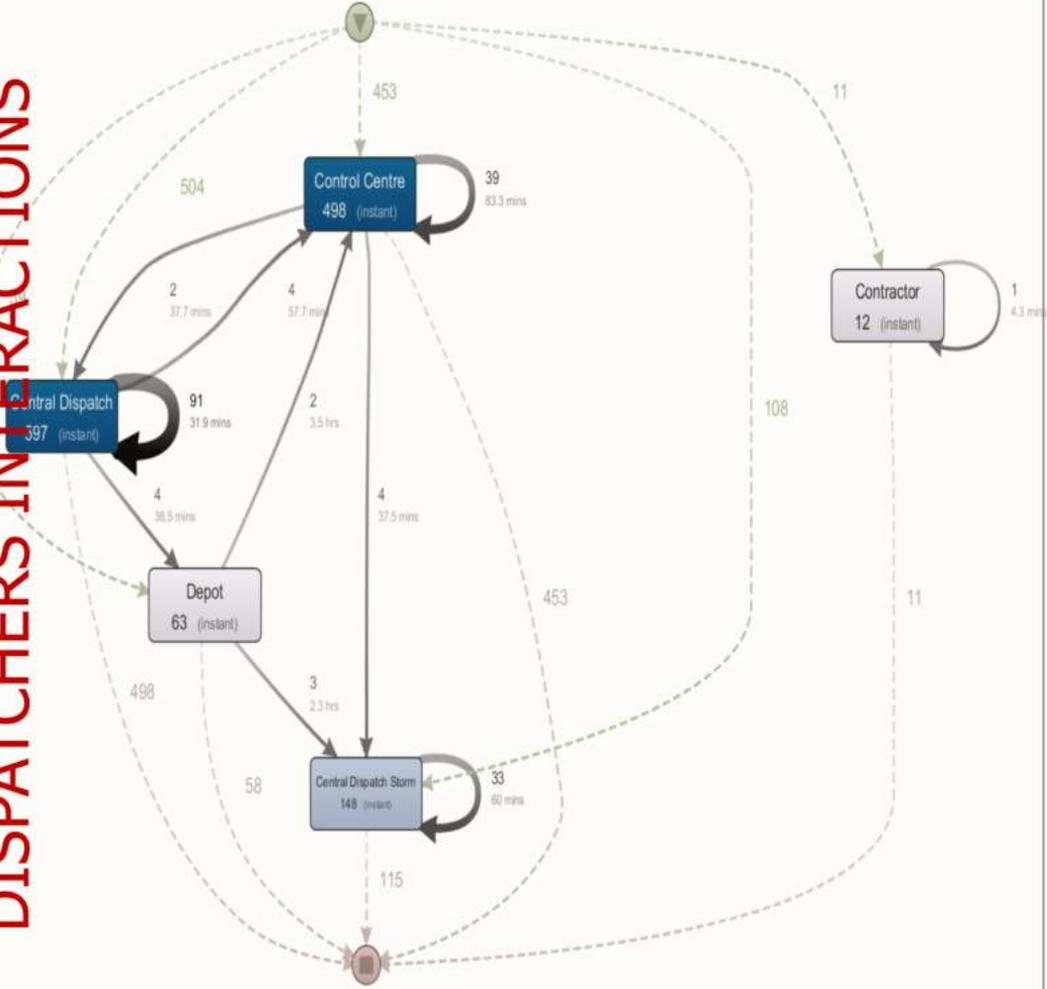


# Resource behaviour analysis

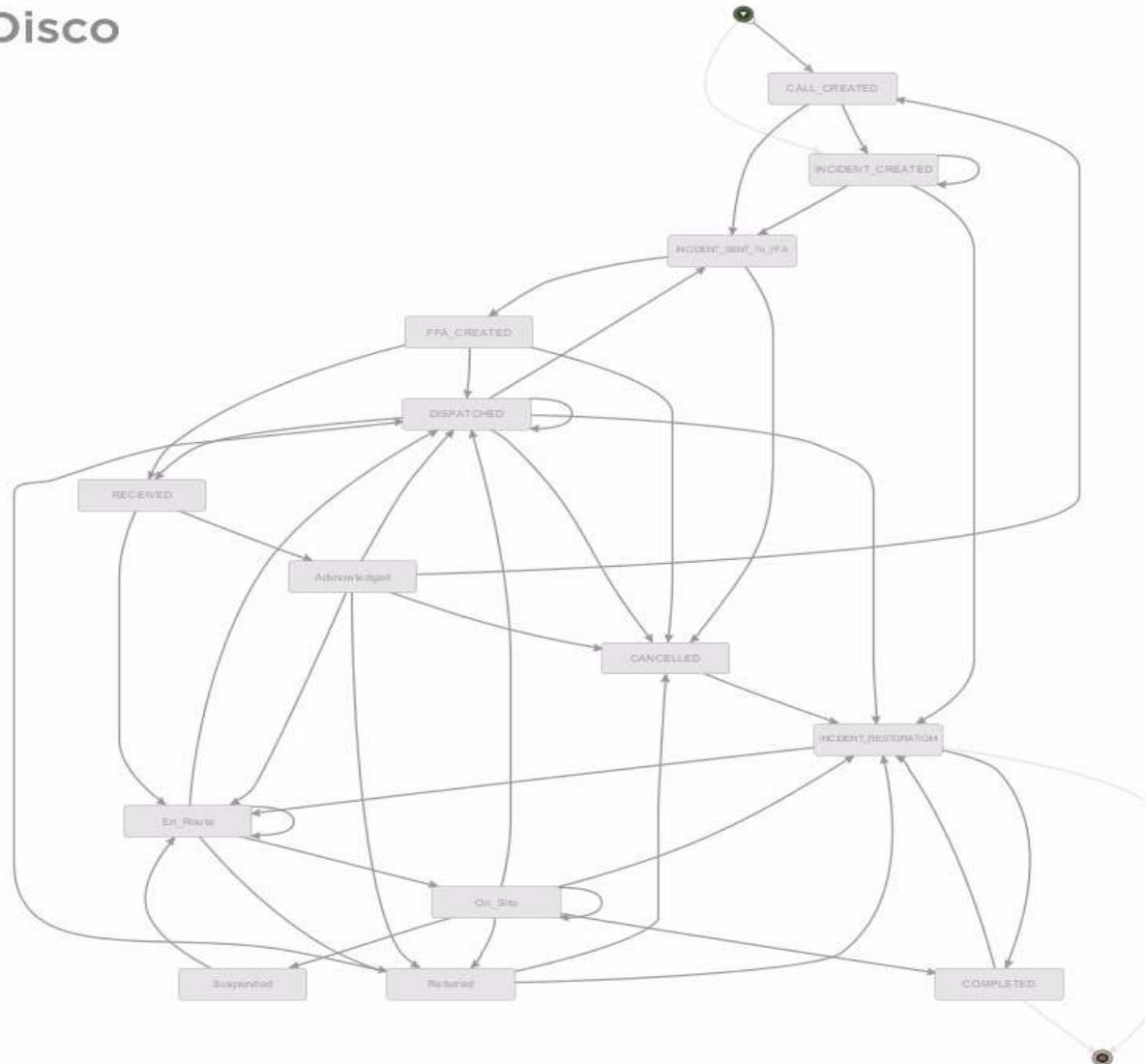
## ROLES INTERACTIONS



## DISPATCHERS INTERACTIONS



# Process Animation



# Summary

---

1. How does the discovered process actually look like comparing to the known business process?
  - Majority of orders from the discovered process follow the same path as the known business process (11 events/activities)
  - Most active cases occurred during business hours
2. Do we meet the performance targets?
  - Just over half of orders were completed in under 5 hours
  - 45 percent of orders fall in long duration to complete
3. Are there deviations from the prescribed process?
  - Orders were cancelled during En\_Route, On\_Site and Acknowledged activities
  - Majority of Dispatchers cancelled original orders
4. How efficient is the current process?
  - Loops in Dispatched, En\_Route, On\_Site activities.
  - Longest total duration spent during evaluation for orders which were eventually completed.
  - Longest mean duration spent On\_Site before Cancelled activity.
  - From the total Referred orders, the majority were cancelled.
  - There's an obvious idle time for orders dispatched by Central Dispatch Storm group (65 minutes mean duration for 70 orders). Crew Ramp-up (total time) caused this delay/idle time before dispatching.
  - There's a pattern in sending crews to the same area (on the same day) for different priority orders. The Longest duration spent traveling to the same area.

# Recommendations

---

1. How does the discovered process actually look like comparing to the known business process?
  - None.
2. Do we meet the performance targets?
  - KPIs needs to be looked at and maybe changed to address some of the loopholes, idle times according to each role i.e. Dispatcher, Technician etc
  - Long duration orders needs to be looked at and monitored closely
3. Are there deviations from the prescribed process?
  - Some of the Business Rules such as order cancellation during En\_Route, On\_Site and Acknowledged activities needs to be highlighted and shared with the system users
  - Cancellation reasons by Dispatchers could be changed in VSS to enable more accurate reporting and monitoring
  - Training might be needed for Dispatchers to remind them on the cancellation process and the business rules related to this topic
  - Introduce/remind system users of the restricted guidelines on the use of other Dispatchers or Technicians logon detail.
4. How efficient is the current process?
  - Training and KPI measures to minimise loops in (Dispatched, En\_Route, On\_Site) activities
  - Improvement or (process change) might be required to minimise Referred & Cancelled orders. The majority of Referred orders were eventually cancelled
  - Are there any measures/steps that could be implemented to reduce (Crew Ramp-up)?
  - Priority dispatching process needs to be reviewed as currently resending crews to the same area seems costing valuable travel time to the required site.
  - Enable, tune and utilise auto-dispatch in order to efficiently dispatch according to area, priority, skills etc.

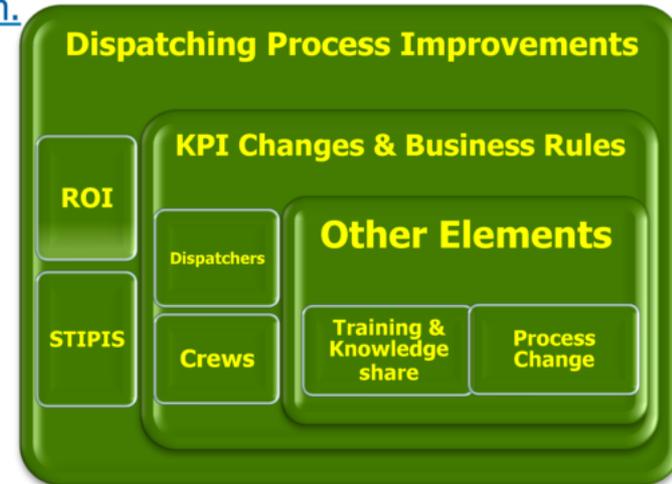
# Improvements and ROI Relationships

The following scenarios are looking at around 1000 order per storm.

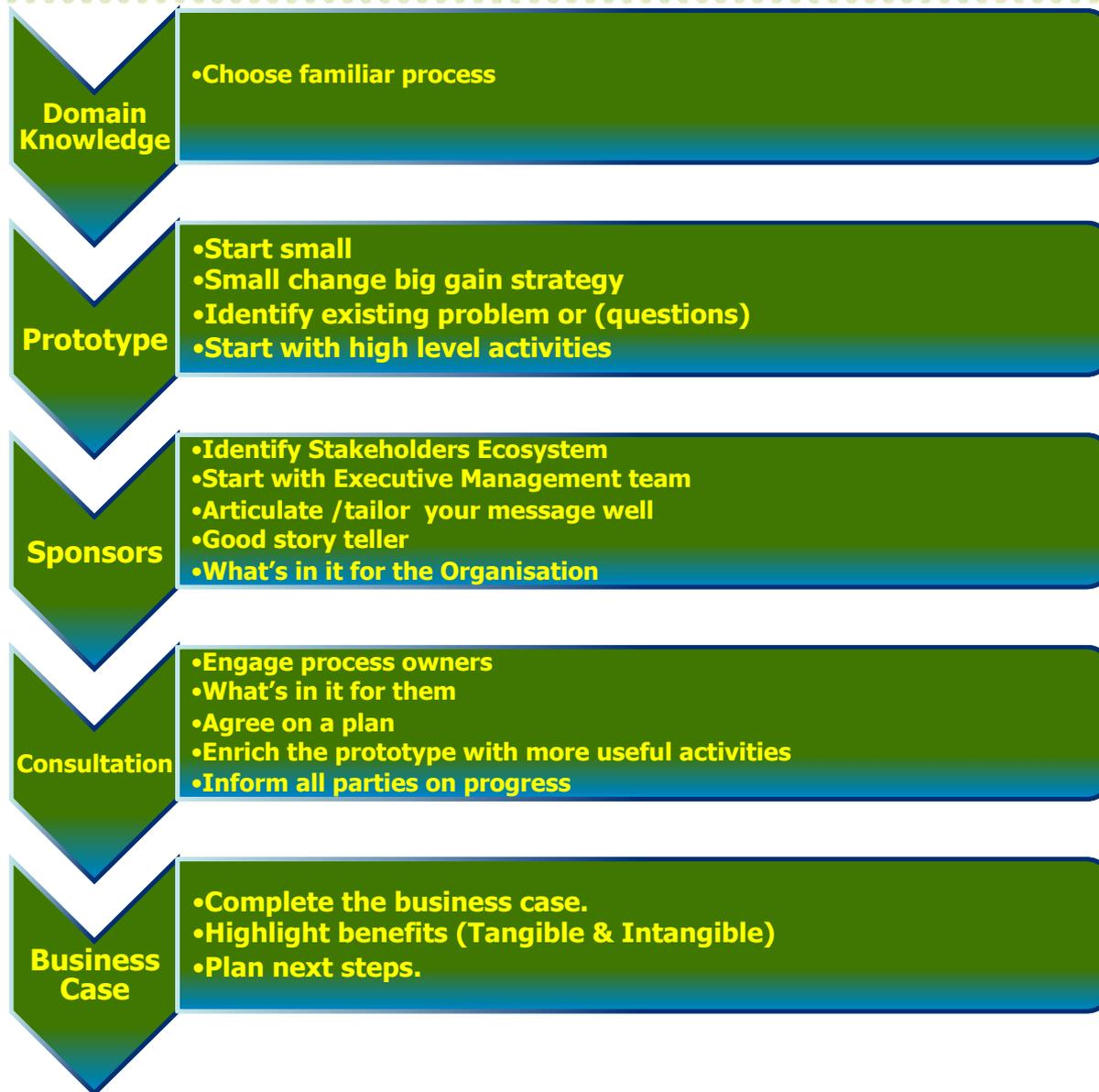
- **Cancelled orders 26%**
  - 26% of the orders were cancelled of this 87% were cancelled by a dispatcher.
  - In total 317 jobs were cancelled assume 50% we can improve on
  - Cost of crews + vehicle = \$500 per hour
  - $\$500 \times 158 = \$79,000$  per storm
  - $\$79,000 \times 13$  (approximately storms p.a.) = \$1,027,000 Benefits.
- **Loops 53%**
  - Dispatch 35%
  - Pulled back and re-dispatched to another crew. Improves dispatch efficiency KPI
  - Crews 18%
  - Double clicking Enroute/Onsite maximises production and minimises idle time
  - Totals target for improvement is 53% or 689 orders
  - Spending 2hrs and reduce to 1hr = 1 hour saving to dispatch and rectify
  - $689 \text{ orders} \times 1 \text{ hr quicker restoration (STIPIS claim \$\$)}$
  - Benefits to be determined
- **Referred 13%**
  - 63% of referred orders were cancelled by EAI and not PowerOn. PowerOn would then automatically recreate (Average duration 3 Hrs)
  - When an order is cancelled not in PowerOn it leads to PON unable to reconcile it with FFA
  - In all cases when unable to reconcile, leads to a GSL payment @ \$100 GSL \* 99 = \$9,900 \* by 13 storms = \$128,700 pa
- Total benefits: **\$1,027,000 (Cancelled) + \$128,700 (Referred) = \$1,155,700**

## □ **Additional benefits (STIPIS)**

- The improvements/efficiency of Dispatching process leads to a quicker restoration time.
- Up to 28 million to claim/savings from the AER (Australian Electricity Regulator) for power restoration on time.

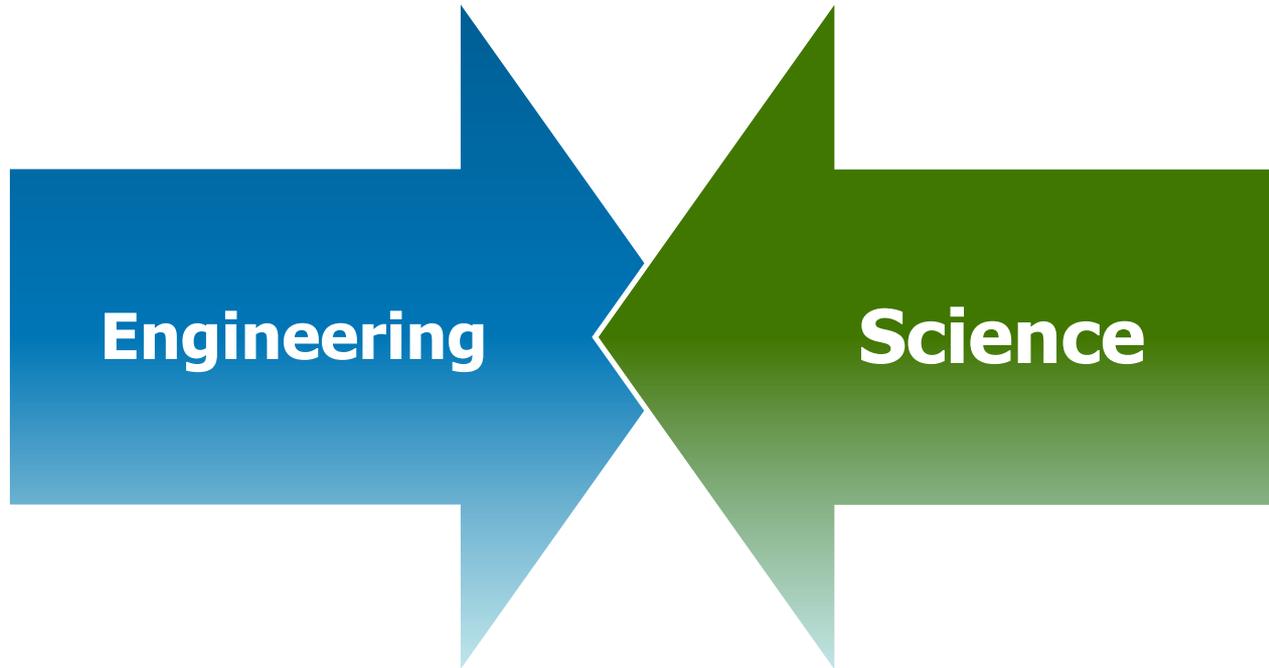


# Data Science Penetration Strategy



# Beware

---



# Q&A..

---

