



A14 Integrated  
Delivery Team



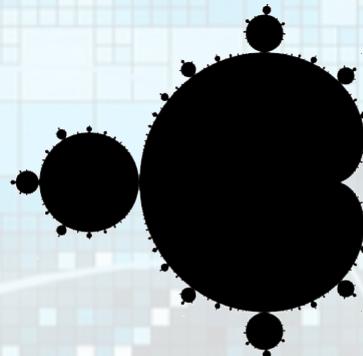
## CHALLENGE #8

# Sentiment Analysis of Observation Data

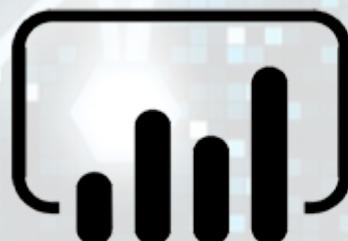
*Users enter comments into our Observation app, but generic tools struggle with construction terms, British English, and sarcasm. Are we able to get better indication of sentiment from the data?*

## Technology used

- Python
  - Machine Learning
- TextBlob
  - Library
  - Common sentiment analysis
  - Input text, output polarity and subjectivity
- PowerBI



TextBlob



Power BI

# Project:Hack5

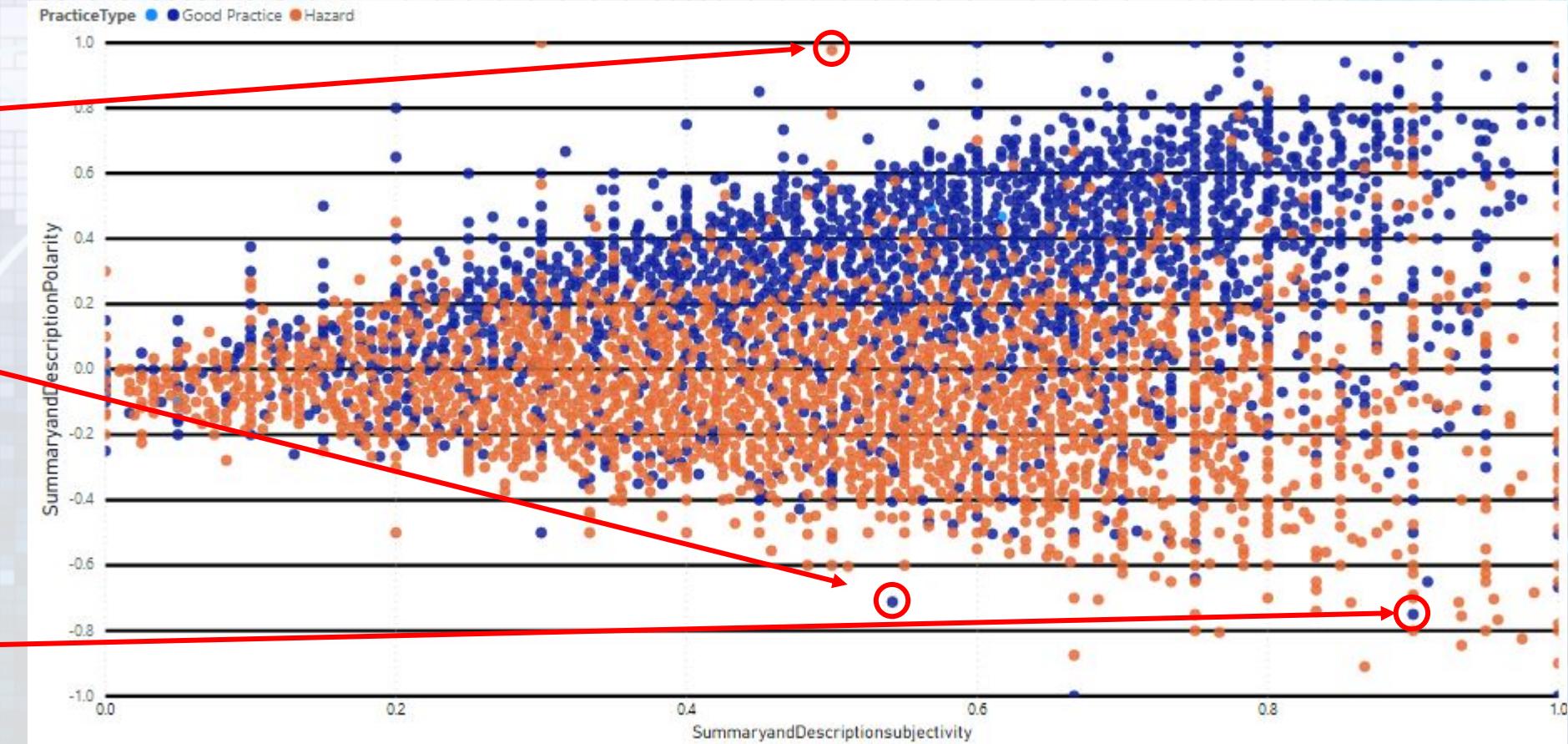
A14 Integrated  
Delivery Team



Too many people using 1 Welfare unit. 11 people! 1 microwave! 30 minute break!

Jordan has been moved from gate to gate but instead of complaining he just got on with it and continued to work hard !!!! Employee of the month  
????

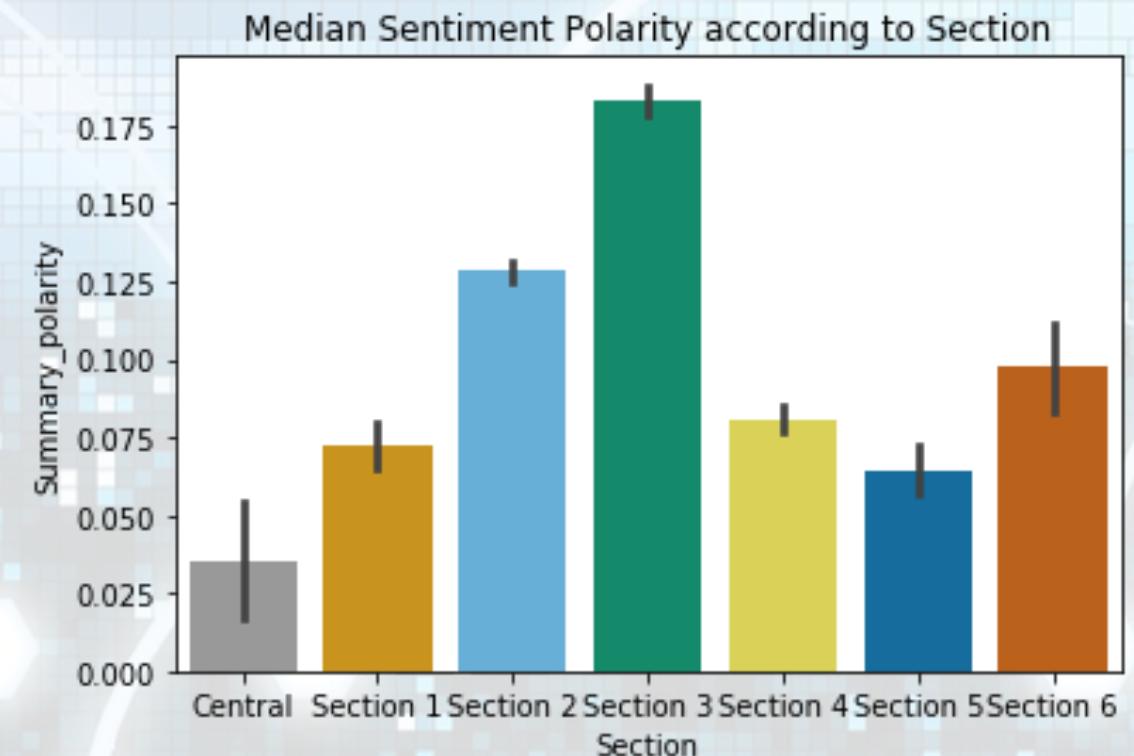
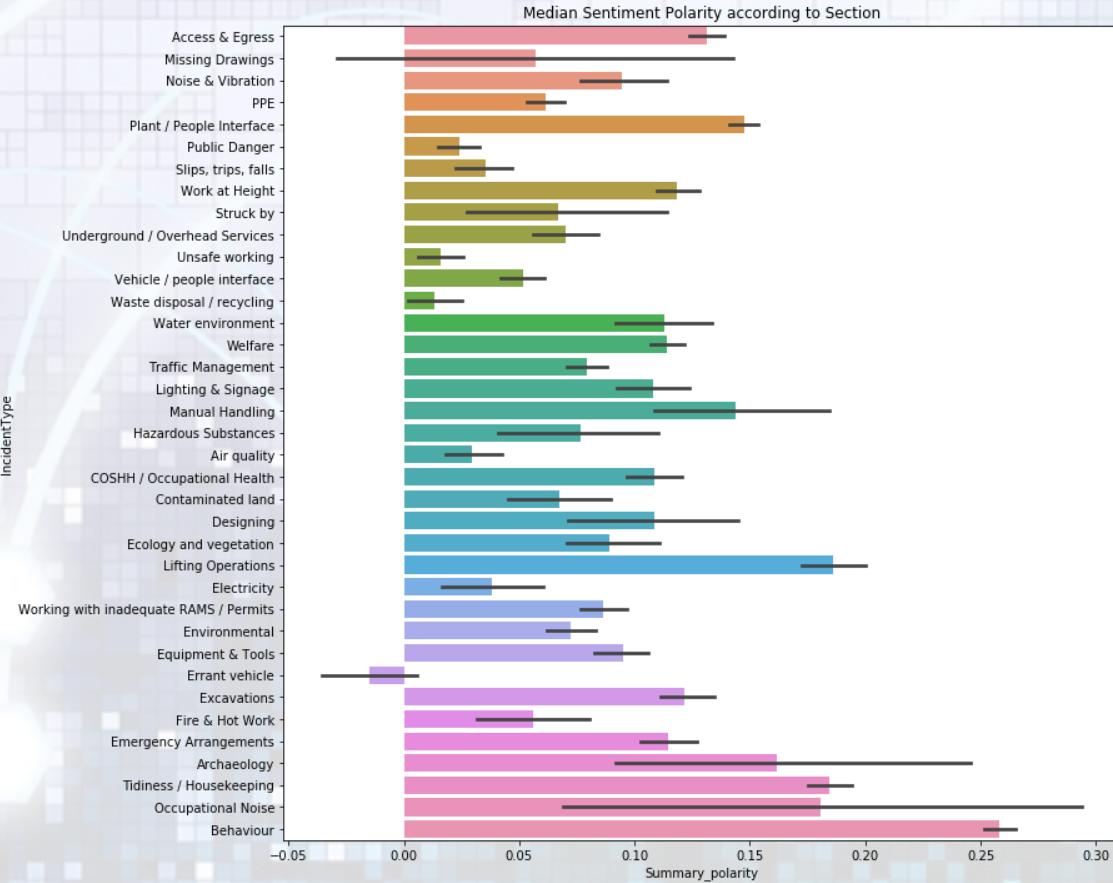
Alan Faulks Roller Driver not using walkways and crossing; cutting corners was spoken to and asked to change behaviour or walk!



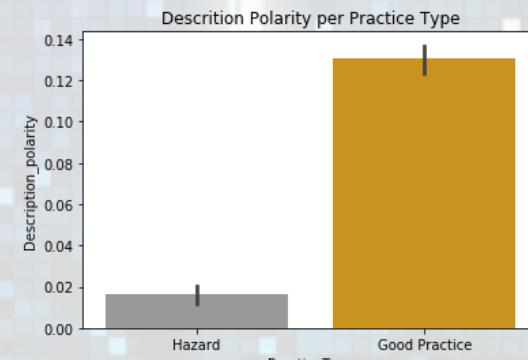
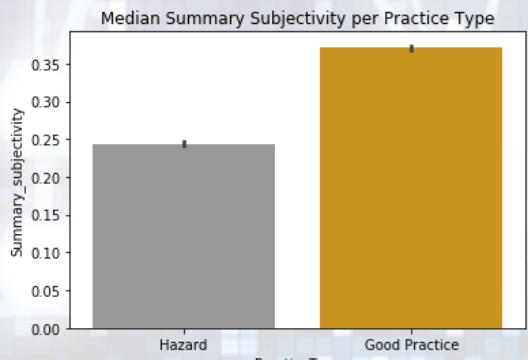
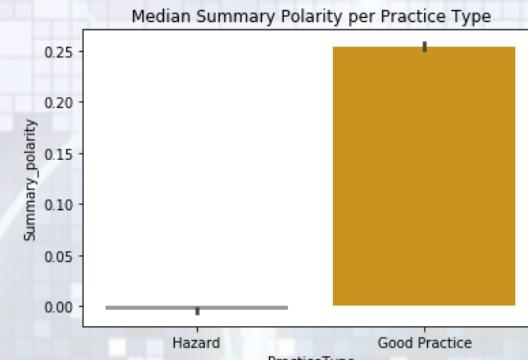
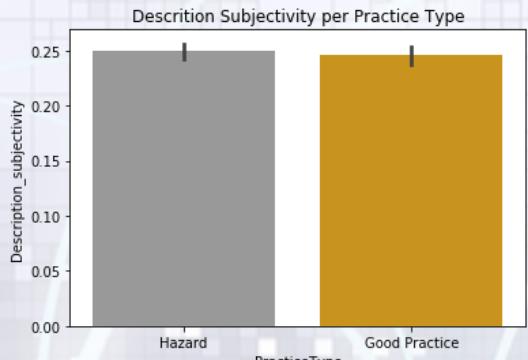
# Project:Hack5

## Project Insights

A14 Integrated  
Delivery Team



## Validating TextBlob Model



## Predictive Analytics

PracticeType	IncidentType	Section	ObservationDateTime	Summary	Description	ActionTaken	HEReportingType	Summary_polarity	Summary_subjectivity
Hazard	Welfare	Section 2	2016-12-09 13:00:00	Cabin was damaged during cleaning process	NaN	None	Hazard	0.00	0.00
Hazard	Working with inadequate RAMS / Permits	Section 2	2016-12-09 13:00:00	RAMS could be better	NaN	RAMS training is in the process of being compl...	Hazard	0.50	0.50
Hazard	Vehicle / people interface	Section 2	2016-12-09 13:00:00	Speeding in borrow pit 1	NaN	???	Near Miss	0.00	0.00

- ‘Practice Type’ predicted by analysis of natural language
- Comparison against actual completed field

Digger jammed  
-0.10

Fell downstairs  
-0.15

Nice one mate, you screwed that one up  
0.60

## Benefits, Innovation & Commercial Application

- Reduce user form filling
- Unstructured data
  - Quantitative score from unstructured data
- Return insights from natural language
- Supports adoption of site technology
  - Most contractors have a similar observation system
  - Potential of voice recognition.
  - Lower training requirements
- People centric
  - Support wellbeing
  - Measure of sentiment as part of day to day activities

## Summary & Going Forward

- Increase data sources
- Feedback to language model
  - Teach it 'construction speak'
  - Use of other language models
- Link between types of hazard and sentiment – does sentiment directly correlate with level of risk?
- How is sentiment linked to other data types:
  - Time of day?
  - Weather?
  - Traffic data?
- Forensic investigation – was an event predictable by previous observation sentiment