



FI•CONSULTING

## Foundations in Fraud Analytics

November 8, 2016

[ficonsulting.com](http://ficonsulting.com)

## About FI

60 client service professionals based in Arlington, VA

Primary service areas include

- Custom Analytics and Research
- Model Advisory and Support
- Application Development and Implementation Support
- Independent Verification & Validation (IV&V)



Data



Analytics



Modeling



Technology



**Nick Heitzman**  
Managing Consultant,  
Data Science



**Jon Hill**  
Senior Consultant



**Kimble McCraw**  
Manager

## Agenda

- Fraud Analytics: Why you should care
- FI's fraud analytics framework
- Data management for Fraud Analytics
- Data analysis for Fraud Analytics

## Fraud Analytics: Why should you care?

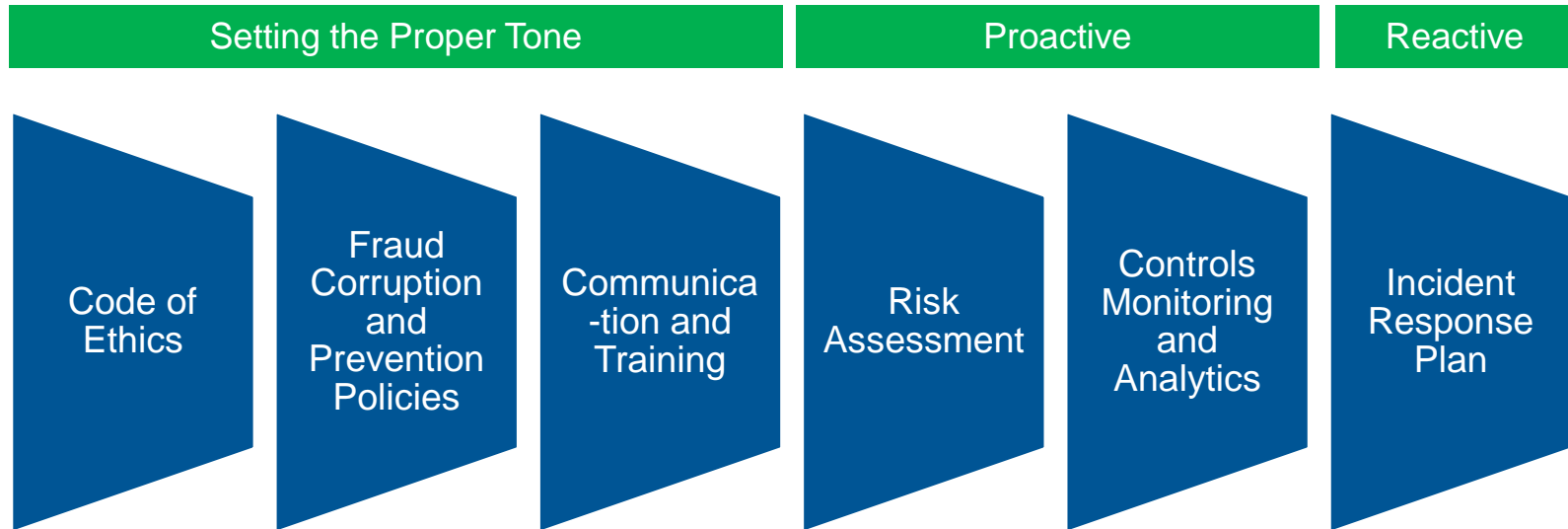
### Case Study: Department of Health and Human Services

The Small Business Jobs Act of 2010 (the Act) required HHS (the Department) to use predictive analytics to Identify improper Medicare fee-for-service claims that providers submit for reimbursement and prevent the payment of such claims.

- In the third year of implementation of the HHS Fraud Prevention System, the system identified:
  - \$133m of recoverable fraud
  - \$453m of potentially un-recoverable fraud
- The HHS OIG determined a 284% ROI on its investment

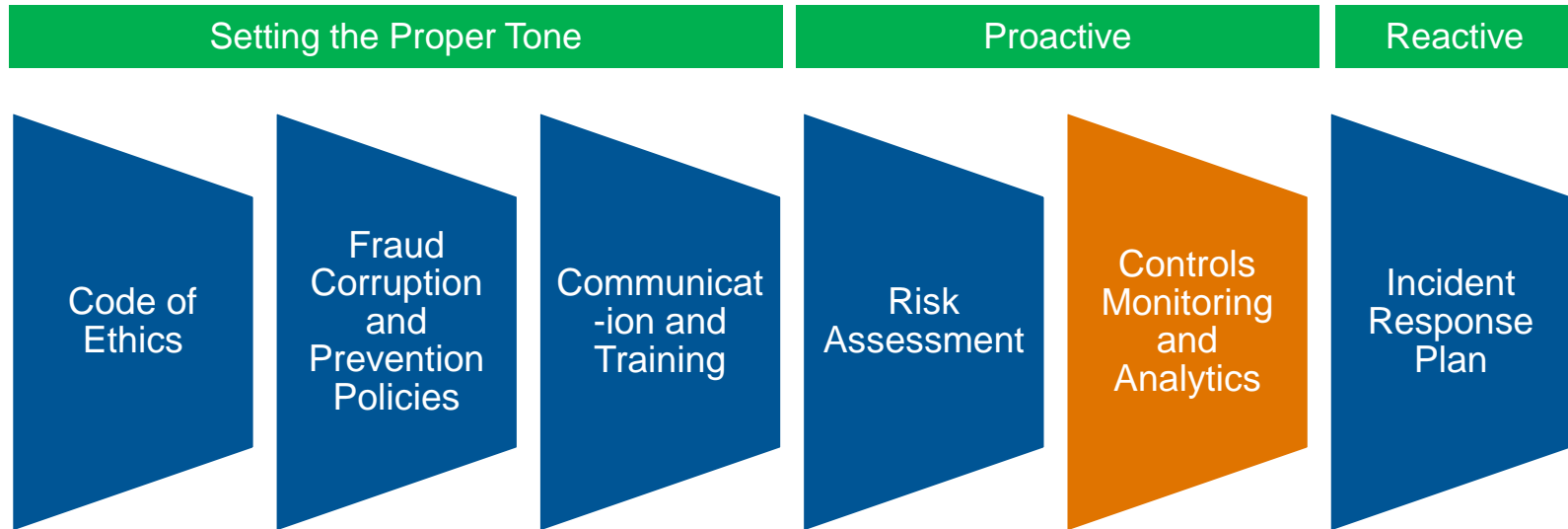
## Lifecycle of Fraud prevention

- Full-scale risk assessment
- Review of policies and procedures
- Rules are created to evaluate whether policies and procedures are being followed



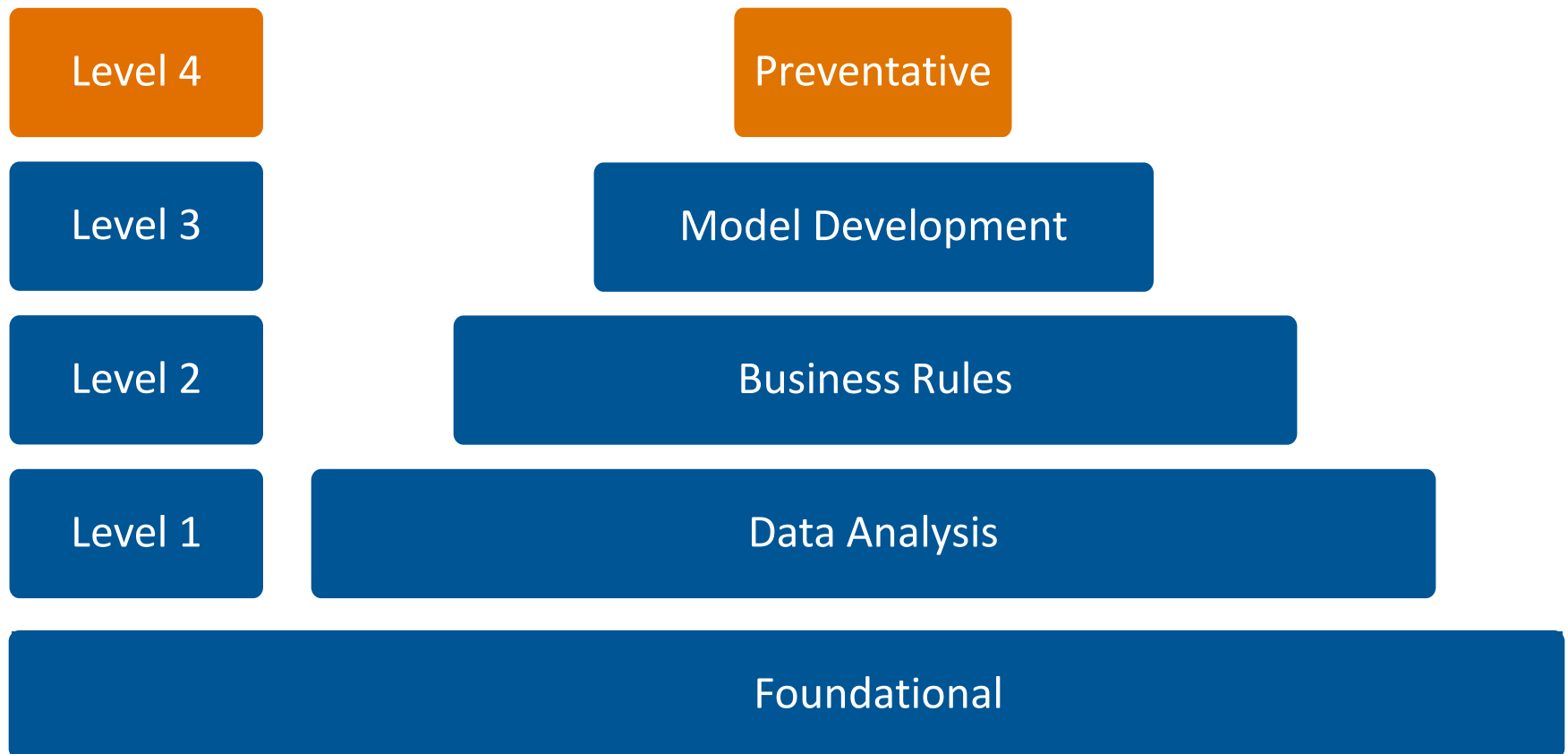
## Lifecycle of Fraud prevention

- Full-scale risk assessment
- Review of policies and procedures
- Rules are created to evaluate whether policies and procedures are being followed



**FI's approach focuses on analytics**

# FI Fraud Data Analytics Approach





# FI Fraud Data Analytics Approach

Foundational

# FI Fraud Data Analytics Approach

Level 1

Data Analysis

Foundational

# FI Fraud Data Analytics Approach

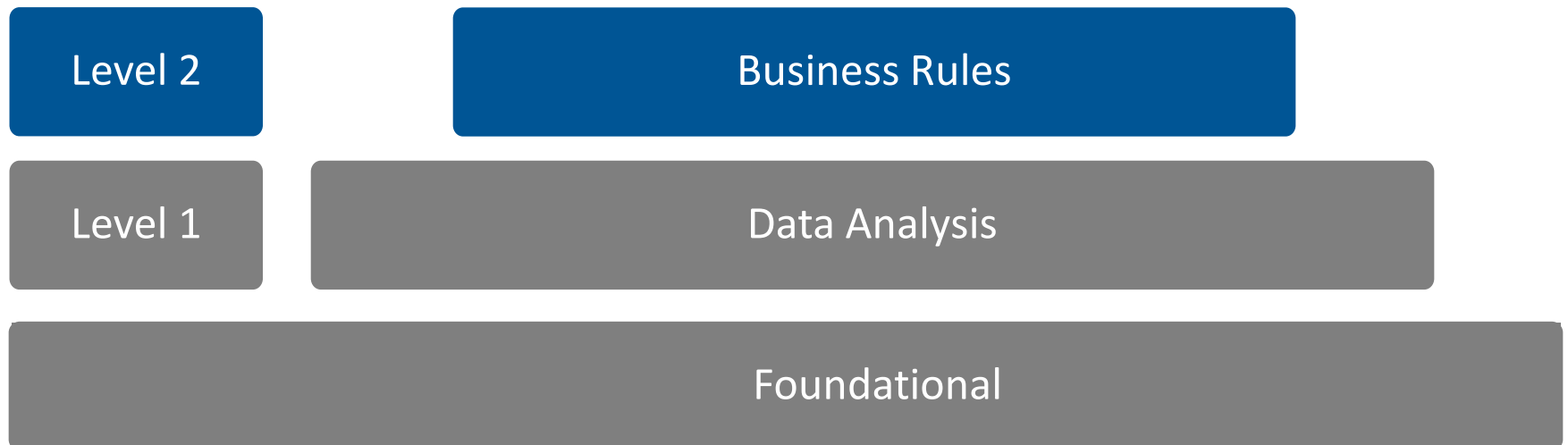
The first step is to examine your data

Level 1

Data Analysis

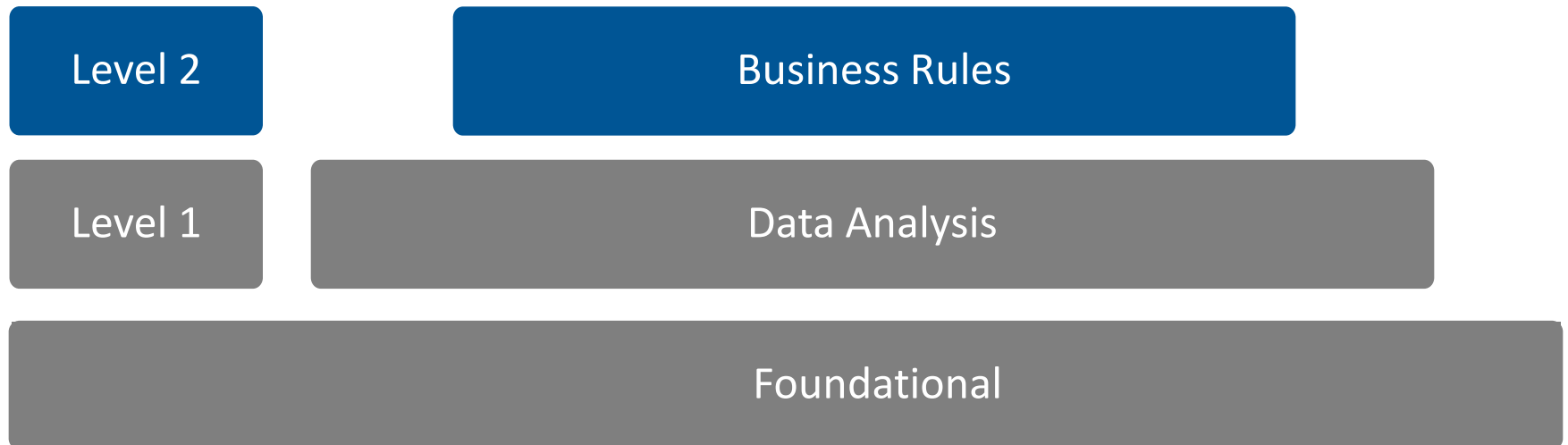
Foundational

# FI Fraud Data Analytics Approach

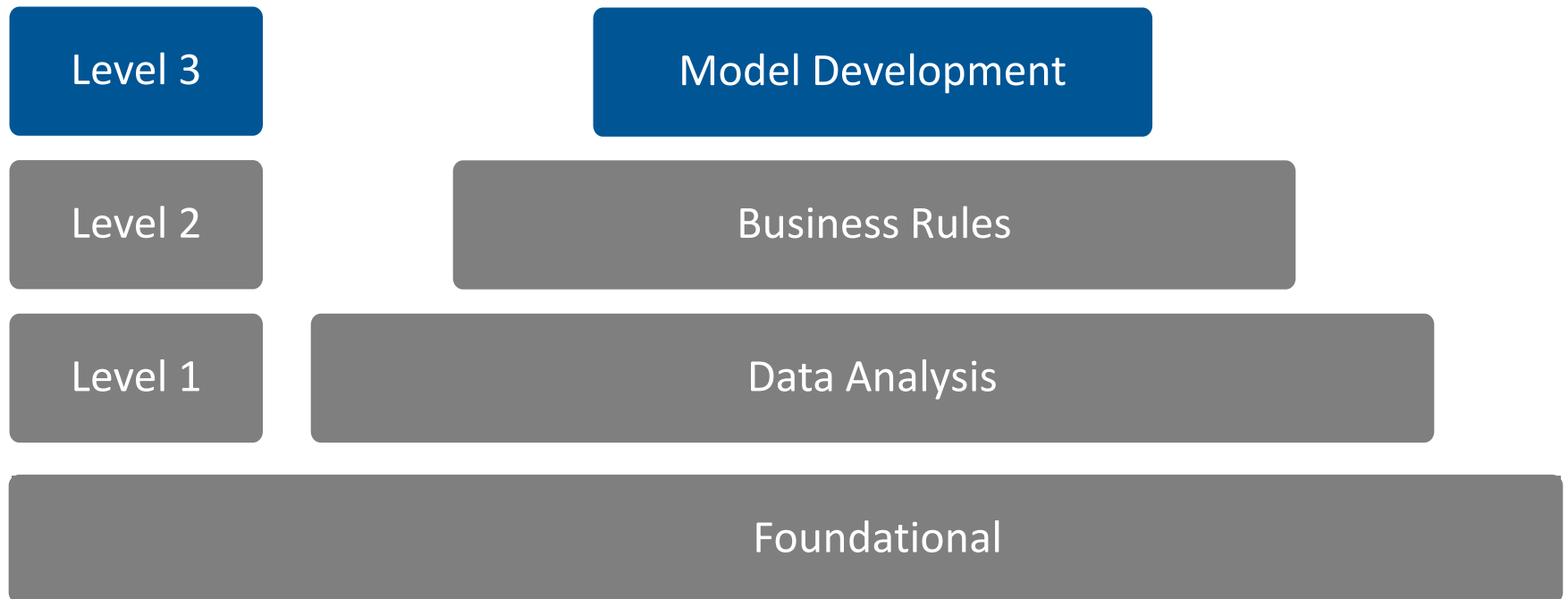


# FI Fraud Data Analytics Approach

Vast majority of Fraud Analytics end at this level

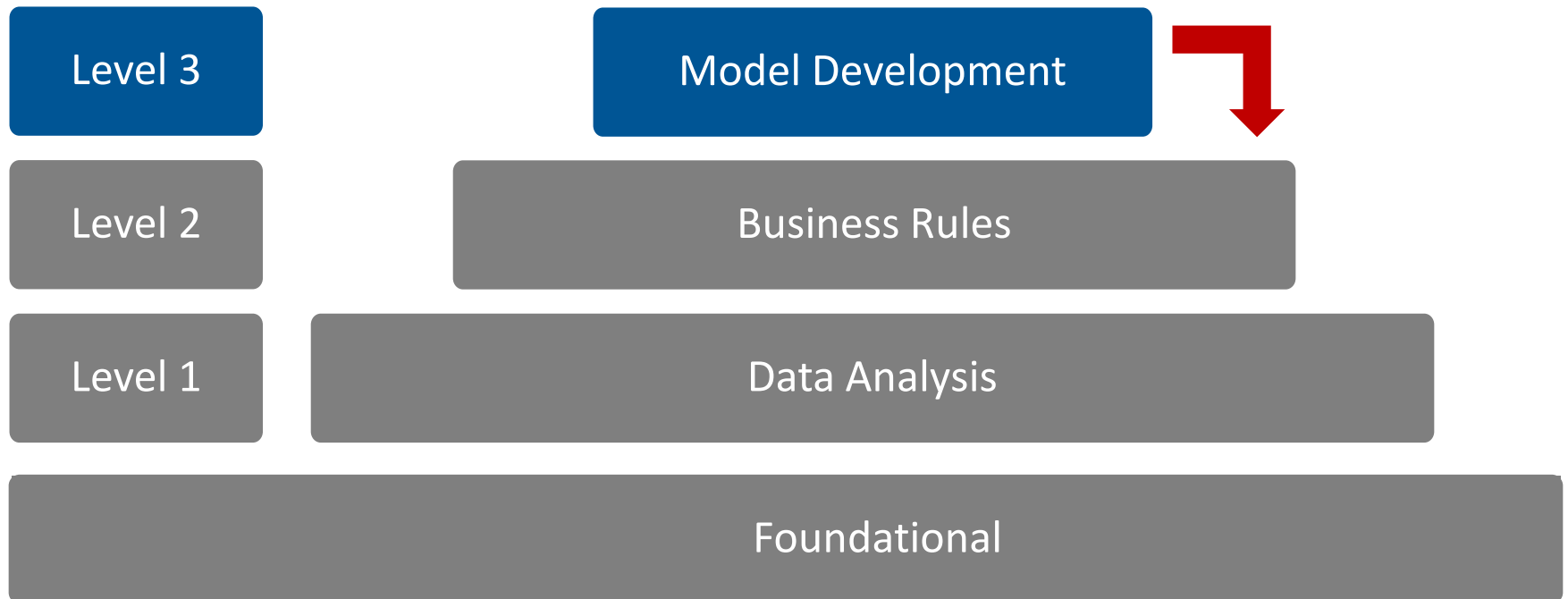


# FI Fraud Data Analytics Approach



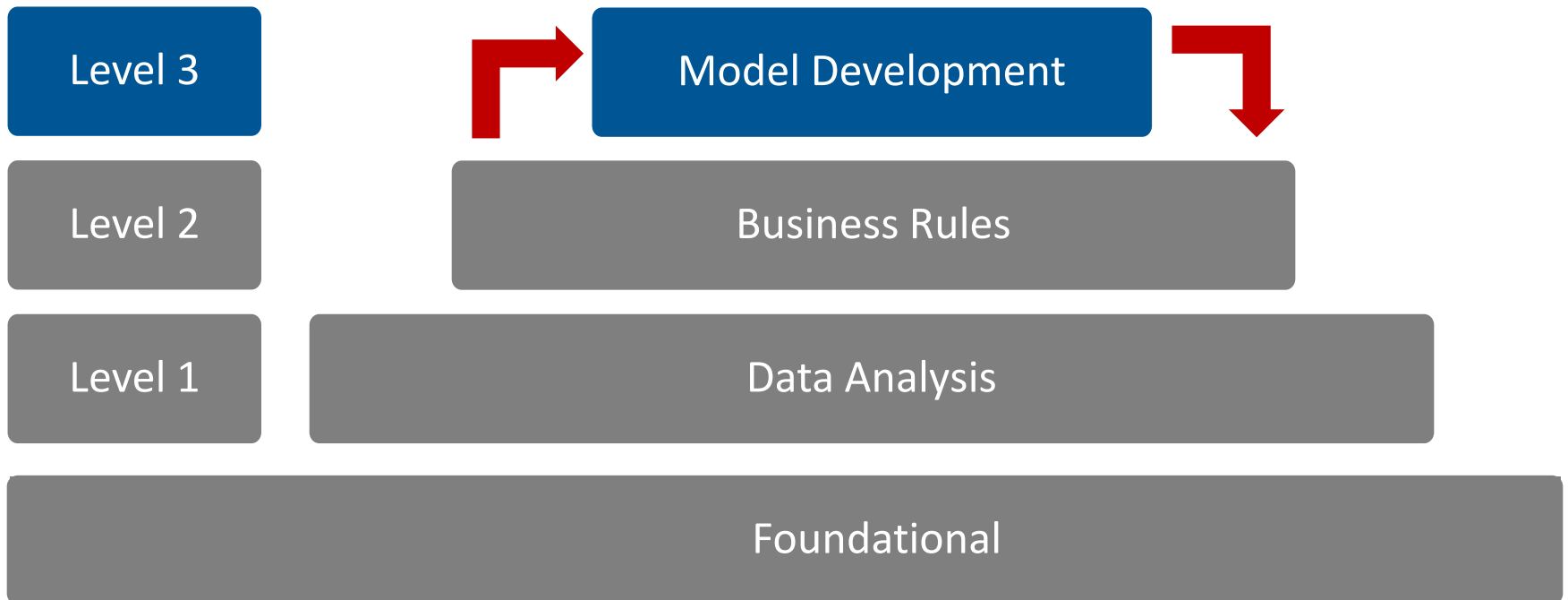
# FI Fraud Data Analytics Approach

These newly identified, high-risk areas help create and refine business rules



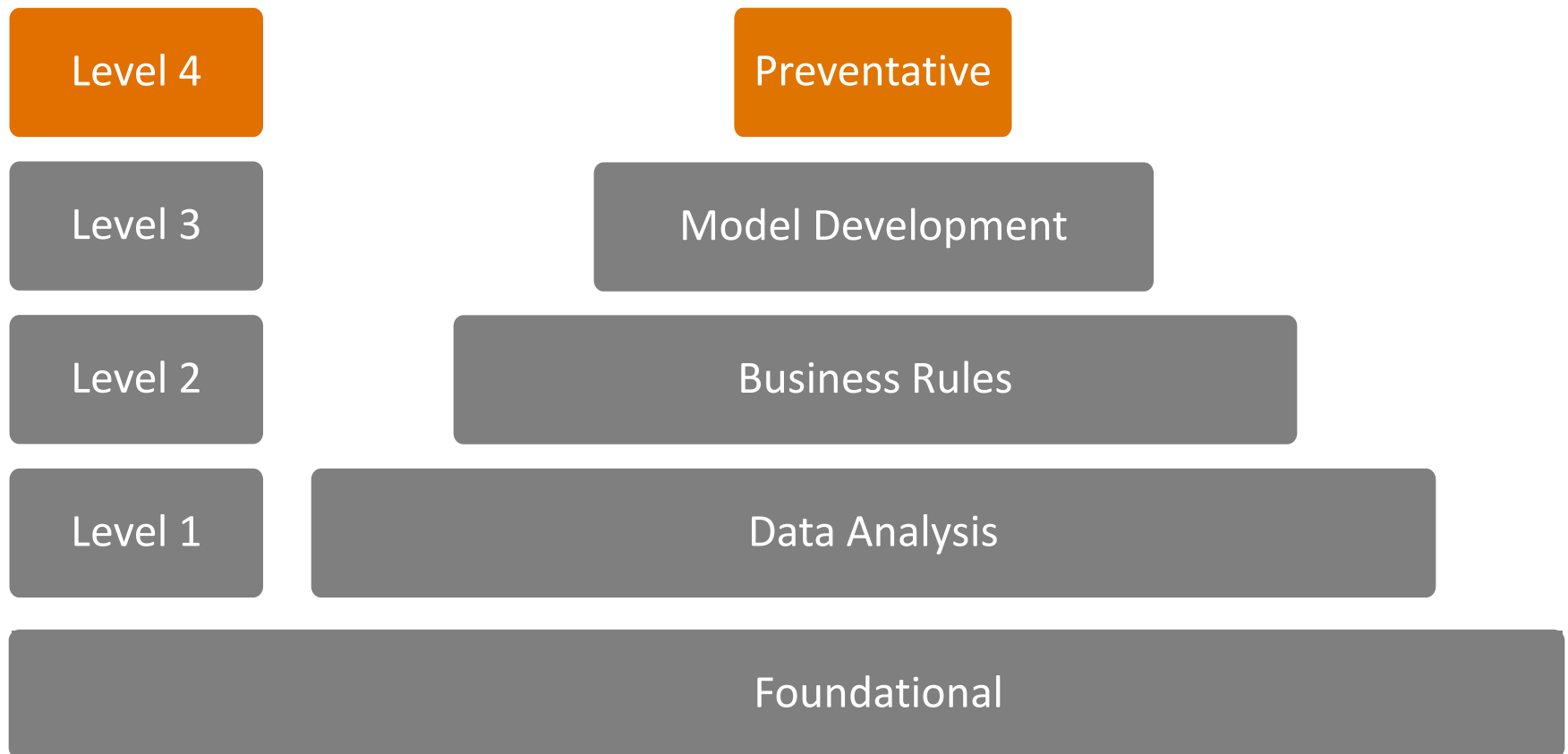
# FI Fraud Data Analytics Approach

Insights from further developed models continuously provide new rules for foundational approach



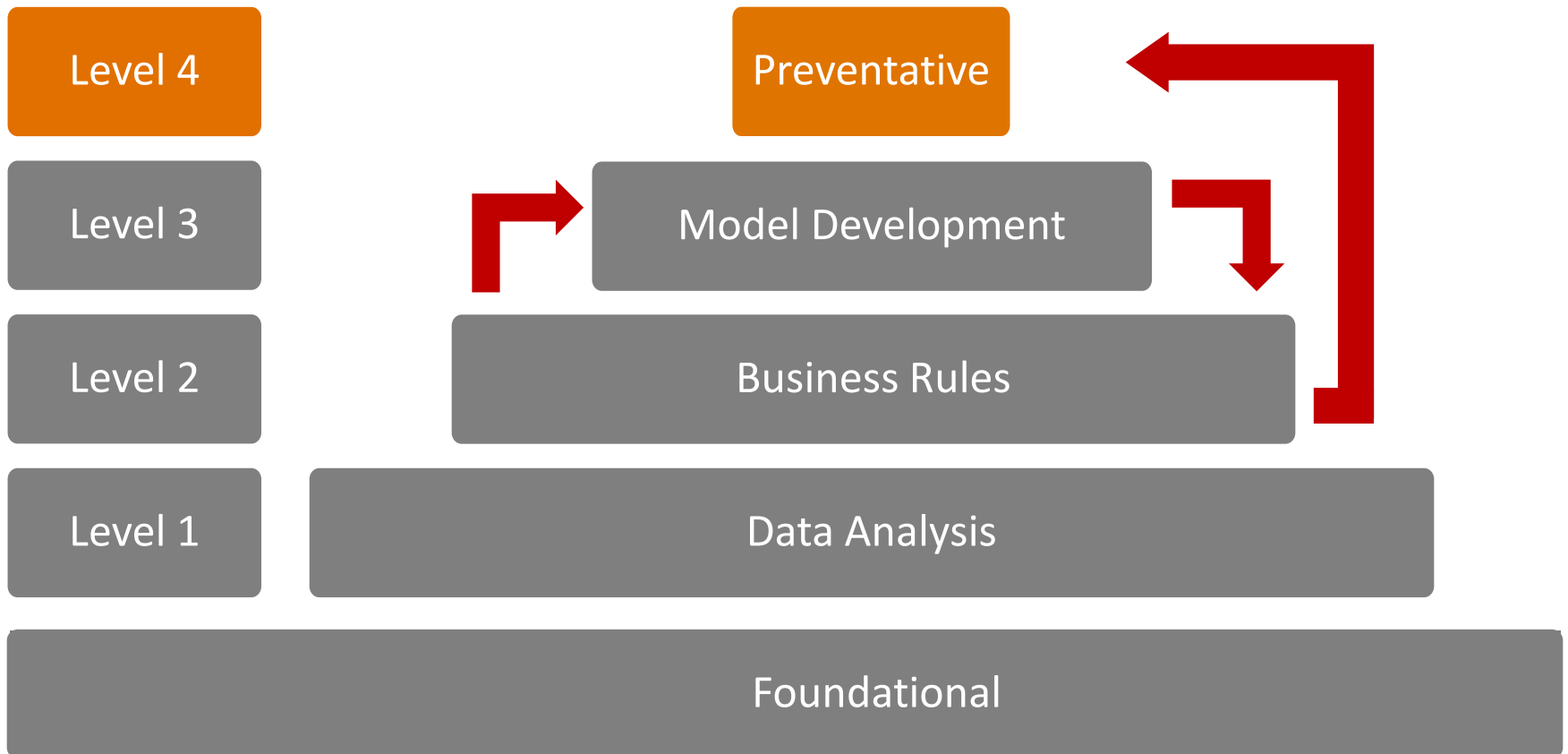


# FI Fraud Data Analytics Approach

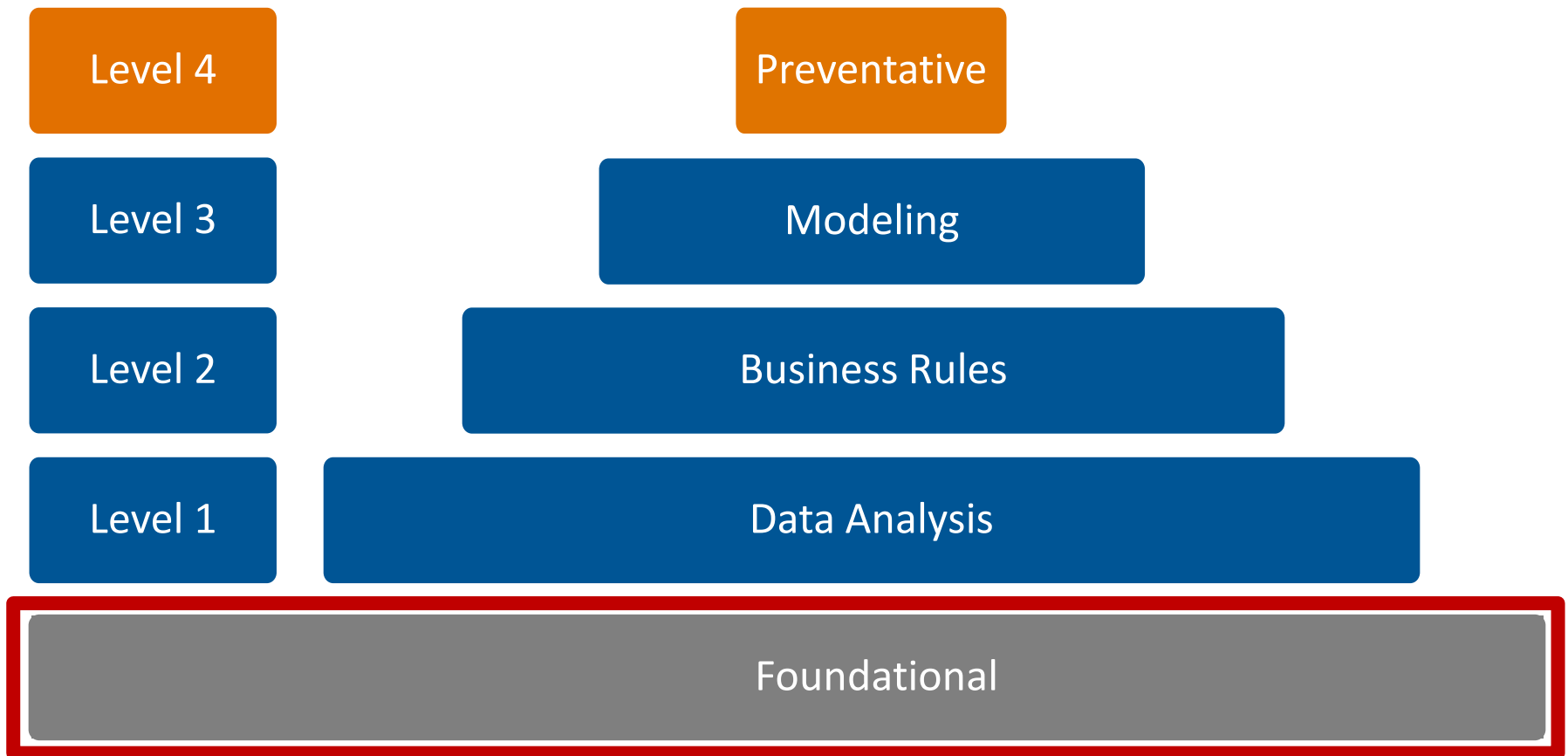


# FI Fraud Data Analytics Approach

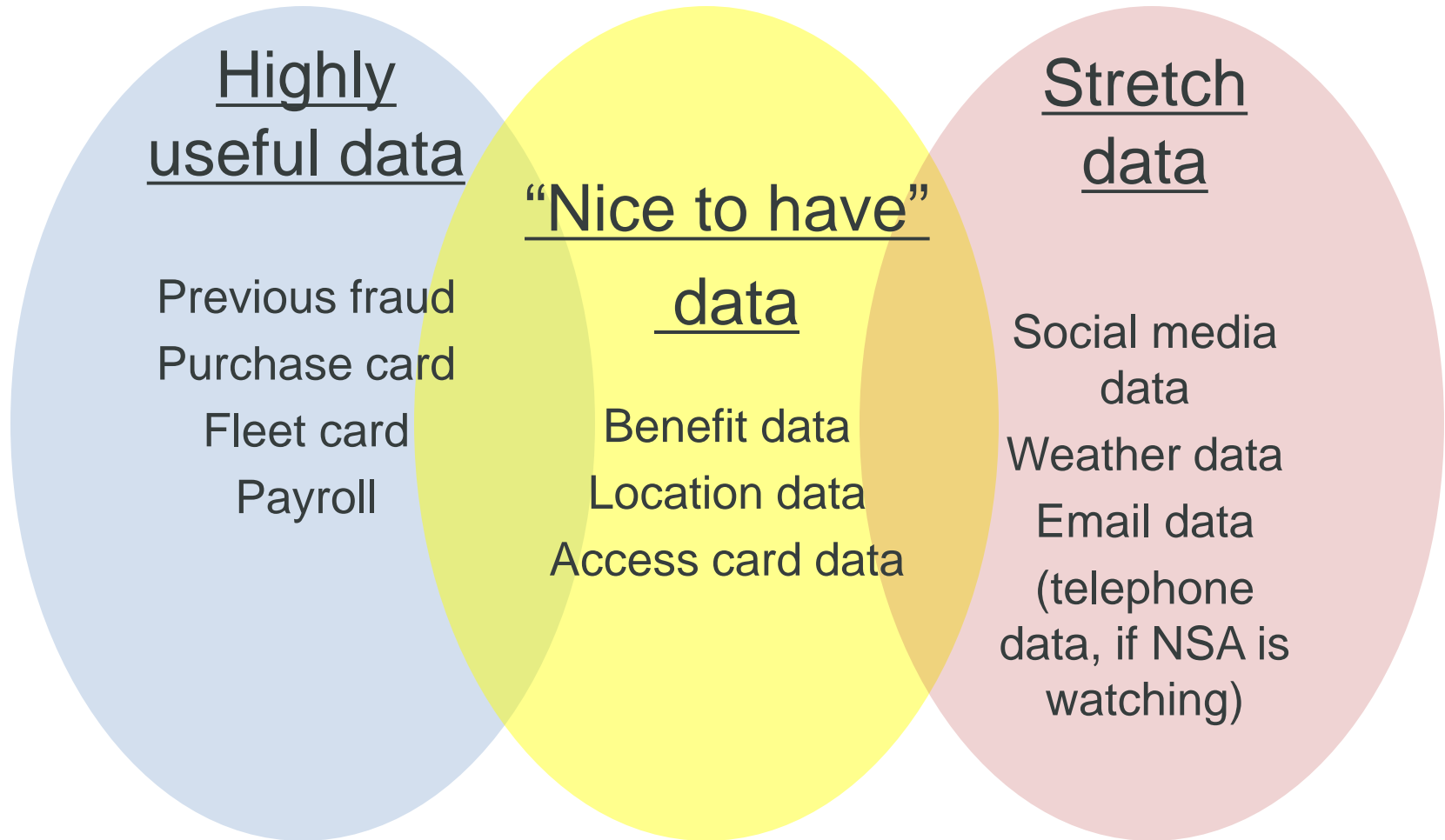
Insights from models provide requirements for future-state, real time preventative monitoring system



# FI Fraud Data Analytics Approach



## What kind of data?



## Data Gathering

### Internal

#### Statutory Data

IPERA, Purchase Card Audit, etc

#### Data Requests

### External

#### API

#### Download

#### Scraping

# Data storage



Oracle



MySQL



SQL Server

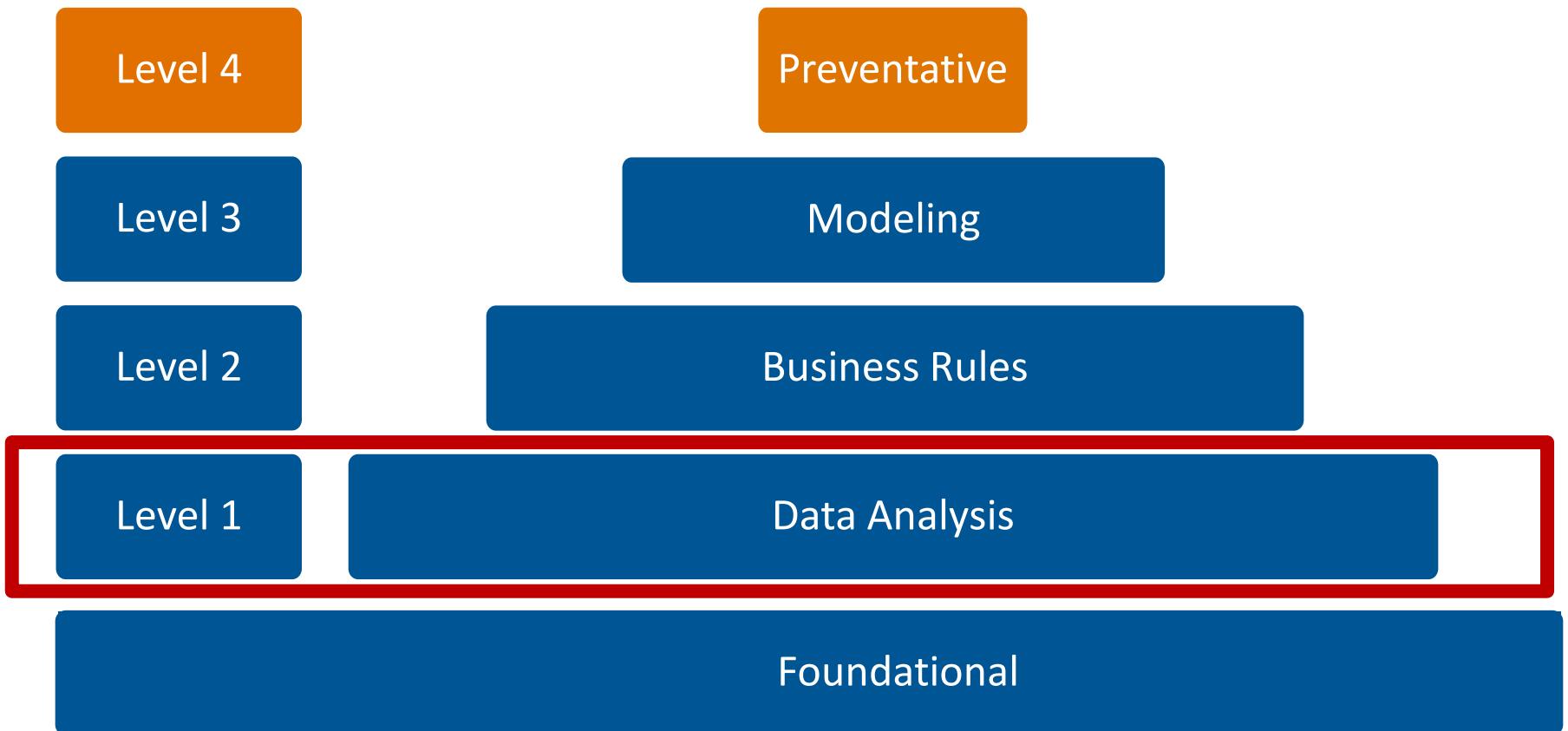
## Case Study: USPS

USPS has many contracts for long distance transportation, which include providing fleet cards.

By adding comparing the geographic terms of contracts with actual fleet card usage locations, USPS identified misuse and contract fraud



# FI Fraud Data Analytics Approach



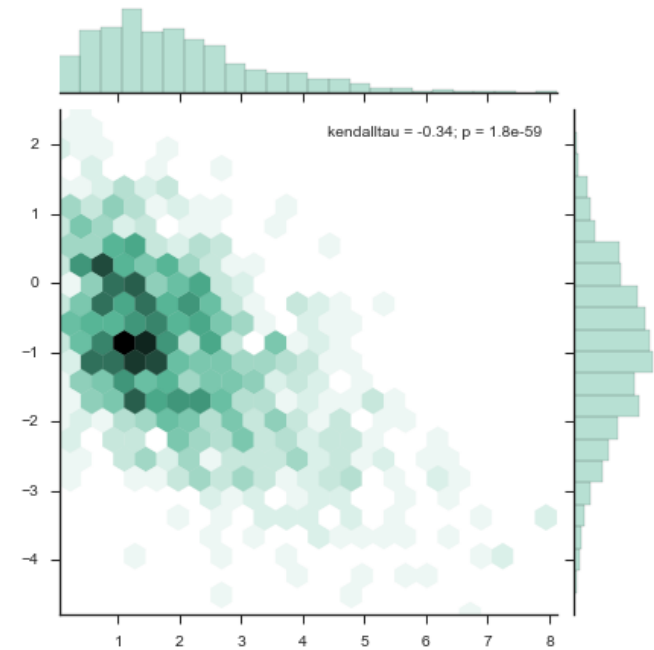
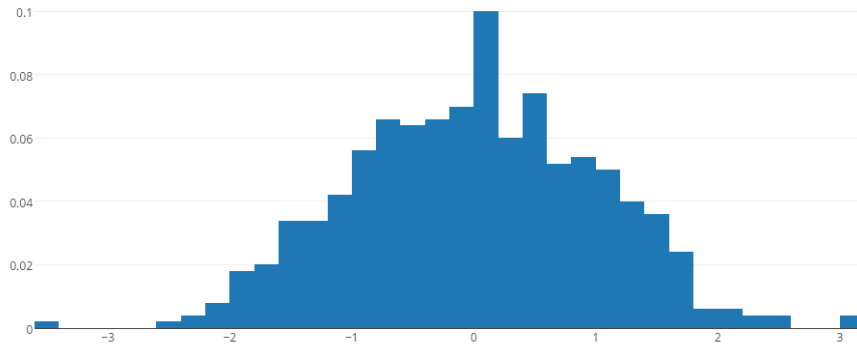


## Benefits of Data Analysis

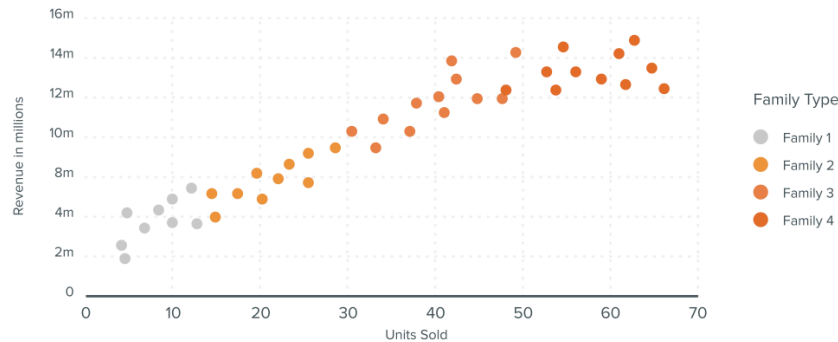
- Identifies “low-hanging fruit” fraud
  - Obvious problems first
- Develops institutional knowledge of dataset
- Drives the creation of business rules
- First step in model specification
- Helps identify the appropriate predictive method or combination of methods

# Fraud Analysis Techniques

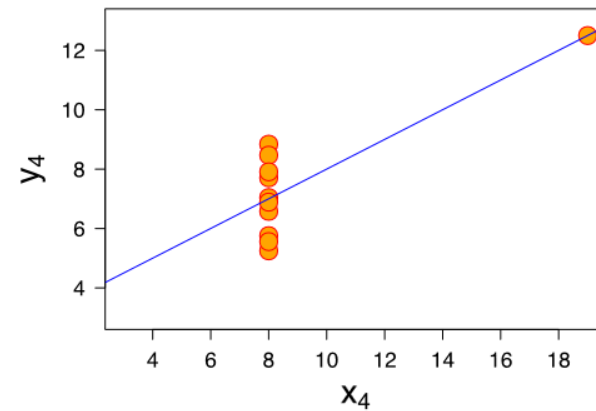
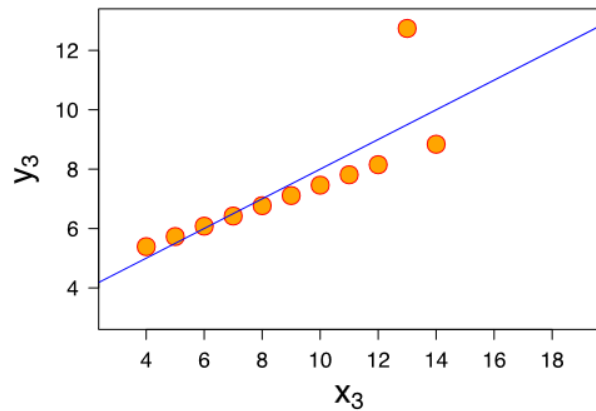
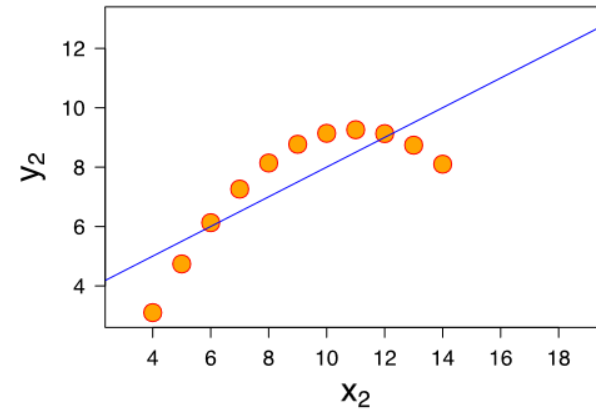
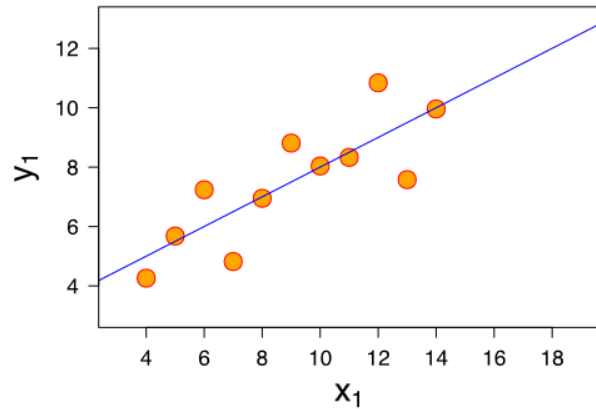
## Visual Analysis



REVENUE, BY PRODUCT FAMILY



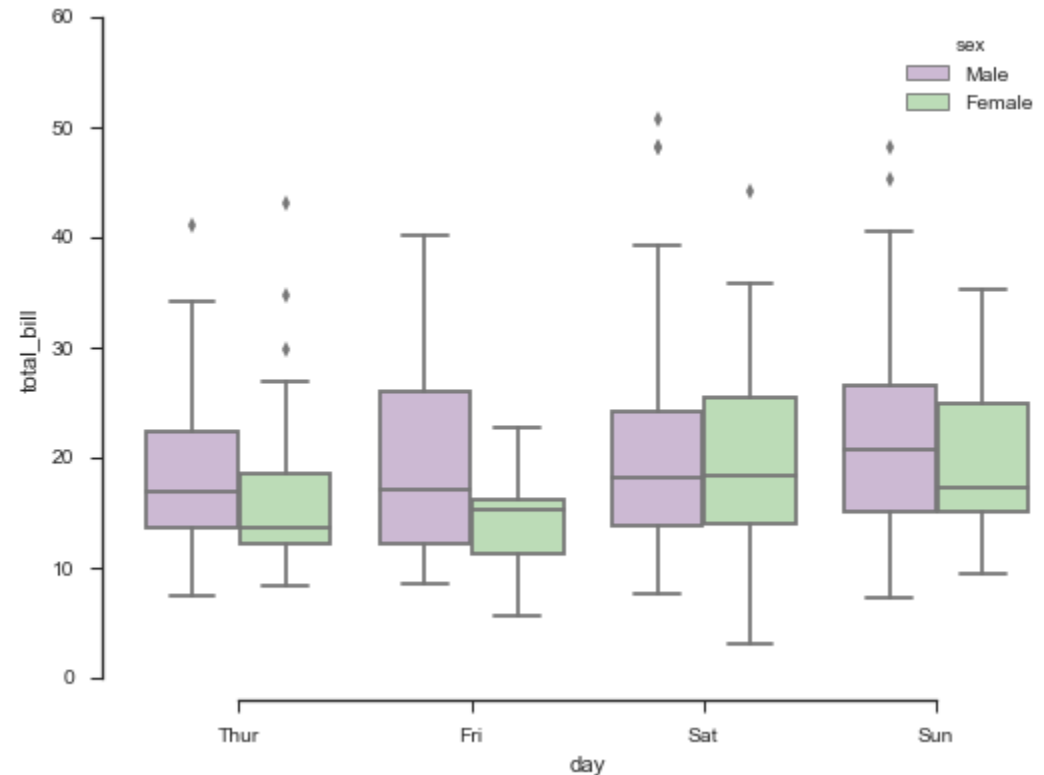
# Anscombe's Quartet



# Fraud Analysis Techniques

## Outlier Analysis

- Box plots can calculate outliers as  $1.5 \times \text{IQR}$
- Z-scores
- Expert knowledge to set cutoffs



# Fraud Analysis Techniques

## Benford's Law

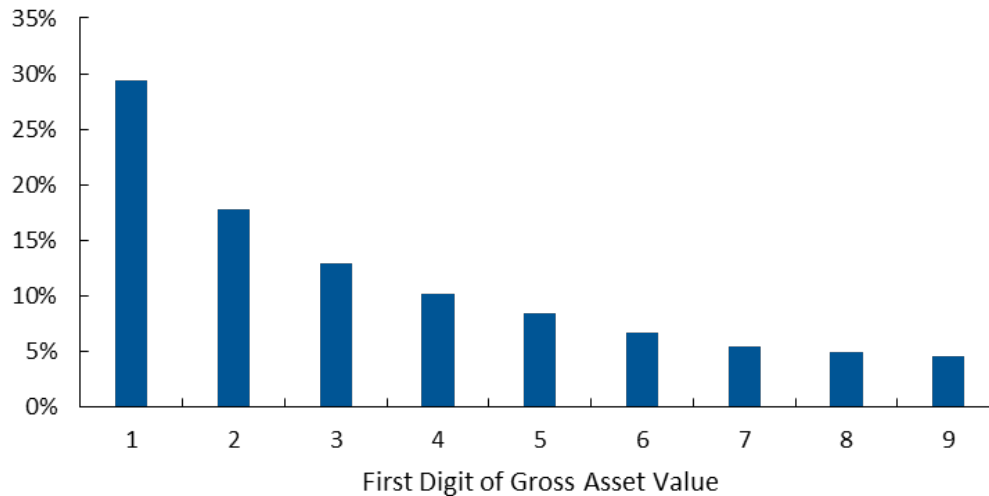
Phenomenological law predicts the distribution of the first digit of many sets of naturally occurring numbers

First Digit	Probability
1	31.5%
2	18.5%
3	13.1%
4	10.2%
5	8.3%
6	7.0%
7	6.1%
8	5.4%
9	4.8%

# Fraud Analysis Techniques

## Benford's Law

Actual Application of Benford's Law

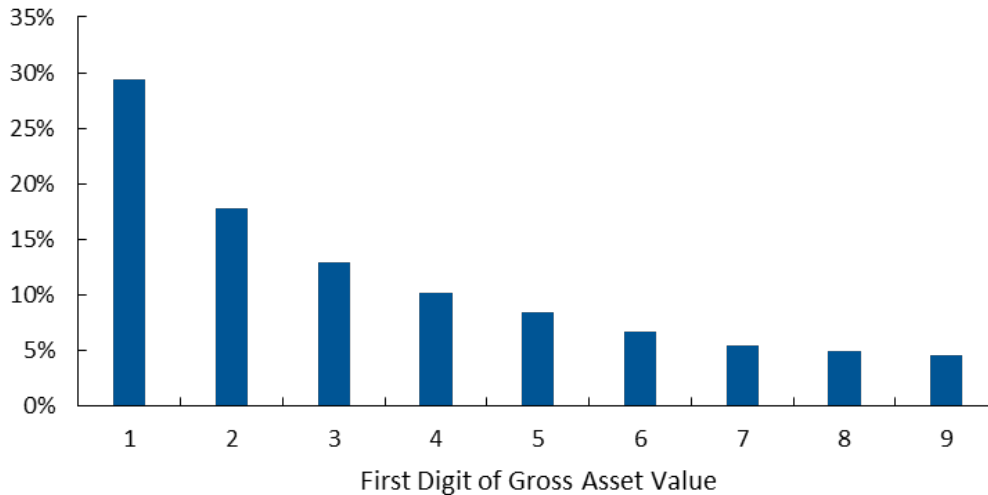


First Digit	Probability
1	31.5%
2	18.5%
3	13.1%
4	10.2%
5	8.3%
6	7.0%
7	6.1%
8	5.4%
9	4.8%

# Fraud Analysis Techniques

## Benford's Law

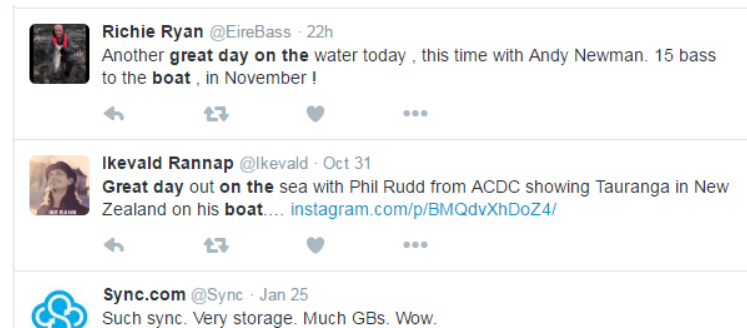
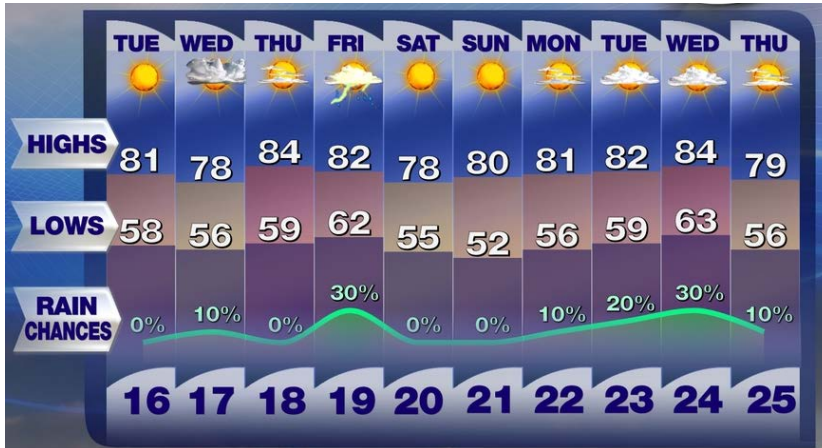
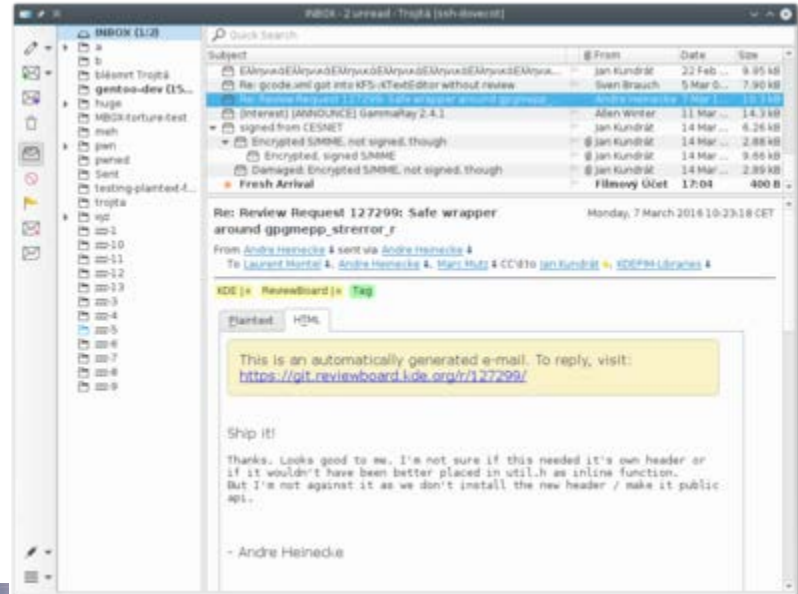
Actual Application of Benford's Law



First Digit	Probability
1	31.5%
2	18.5%
3	13.1%
4	10.2%
5	8.3%
6	7.0%
7	6.1%
8	5.4%
9	4.8%

In a fraud analytics scenario, distributions remarkably different than this would indicate either bad data or potentially fraudulent activity

# Where things could go





“Data problems should be iterative. Start simple. Solve a small problem. Explore the data. Then solve a harder problem. Then a harder one. Each time you take a step, you’ll get ideas on where to go next, and you also get something out at each step. Too many people start trying to solve the entire problem at the beginning, flailing for a long time, usually to discover that it was the wrong problem to solve when they finally struggle to completion. Start with easier problems, learn where to go, and you might be surprised by all the goodies you find along the way.” – Greg Linden

## Contact Us

Nick Heitzman – [Heitzman@ficonsulting.com](mailto:Heitzman@ficonsulting.com)

Jon Hill – [Hill@ficonsulting.com](mailto:Hill@ficonsulting.com)

Kimble McCraw – [Mccraw@ficonsulting.com](mailto:Mccraw@ficonsulting.com)

## GSA Schedules

- OASIS Small Business Pool 2
- Professional Services Schedule (PSS);  
GS00F346CA