RUTGERS

Rutgers Business School
Newark and New Brunswick

Rutgers and AICPA Data Analytics Research Initiative

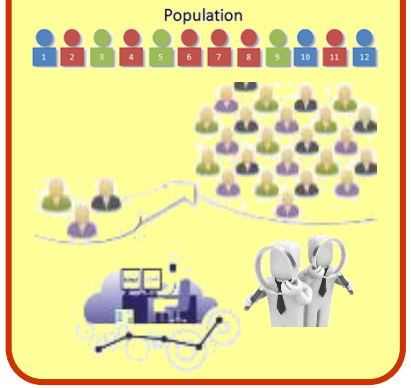
Multidimensional Audit Data Selection (MADS)

PCAOB – CARLab Meeting

BACKGROUND

Traditional sampling approach

New approach

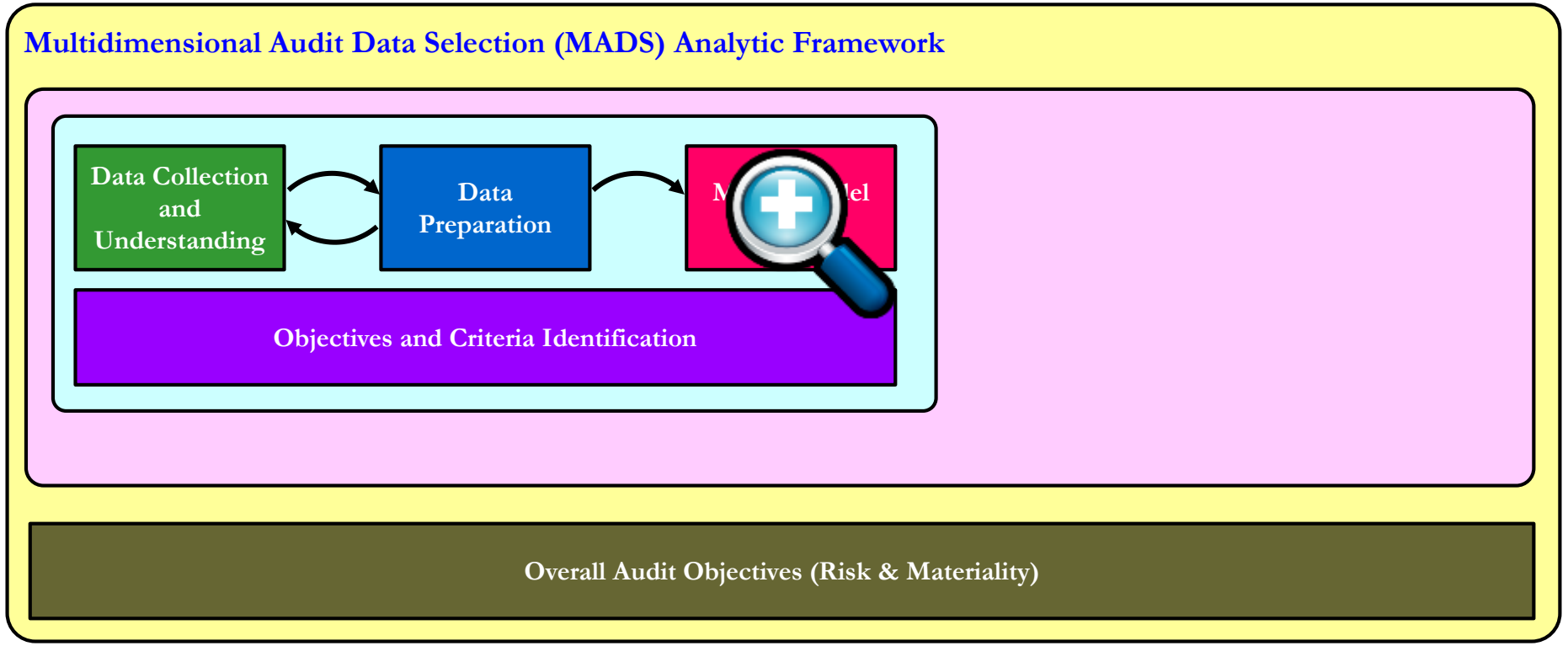


Advance in data processing ability & data analytic techniques allows auditors to evaluate the entire population instead of examining just a chosen sample.

- BUT, often generate large numbers of outliers.
- Impractical for auditors to investigate entire outliers

- Crucial to develop a method that can help auditors effectively deal with large amounts of data, but also assist them to efficiently handle a massive number of outliers.

MADS ANALYTIC FRAMEWORK



- To assist auditors identifying questionable transactions/data in performing substantive test of details
  - Developed based on prior literature and professional guidelines.
  - Modified based on comments from several panel discussions of scholars and auditing professionals.
  - Consist of six components.
- The practice of these six components is guided by the overall objectives of audit, specifically audit risk and materiality.

MADS ANALYTIC FRAMEWORK

RUTGERS
Rutgers Business School
Newark and New Brunswick

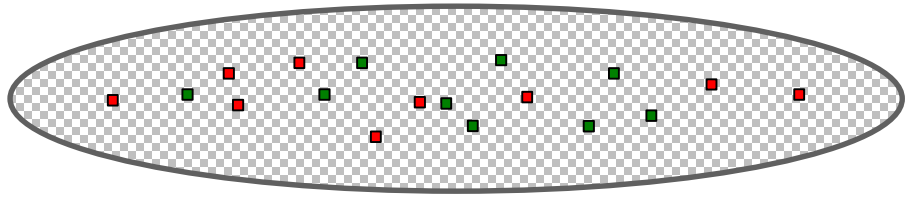
MADS Model Build Process

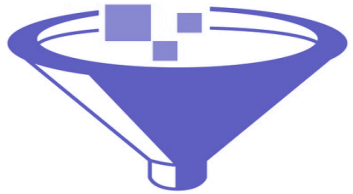
Whole Transaction Data
(Entire Population)

Step 1:
Filters for Significant Potential Risk Factors

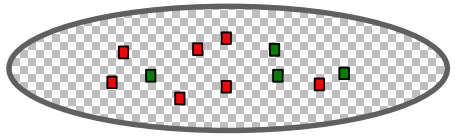
Step 1 Outputs

Step 2:
Data Analytic Techniques

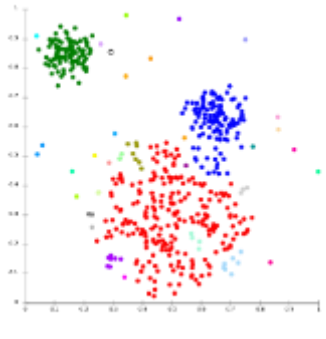


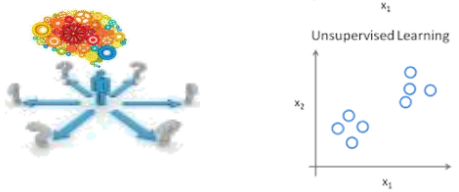


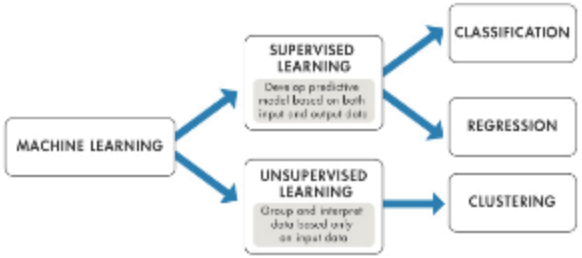
Apply a set of filters to examine significant risks (i.e., What Could Go Wrong) (e.g., duplicate payment)



- Additional Filters
- Visualization Techniques
  (e.g., scatter plots)
- Professional Judgement
  (e.g., knowledge and experiences)
- Outlier Detection Techniques
  (e.g., classification & clustering).







MADS ANALYTIC FRAMEWORK

RUTGERS
Rutgers Business School
Newark and New Brunswick

MADS Model Build Process

Whole Transaction Data
(Entire Population)

Step 1:
Filters for Significant Potential Risk Factors

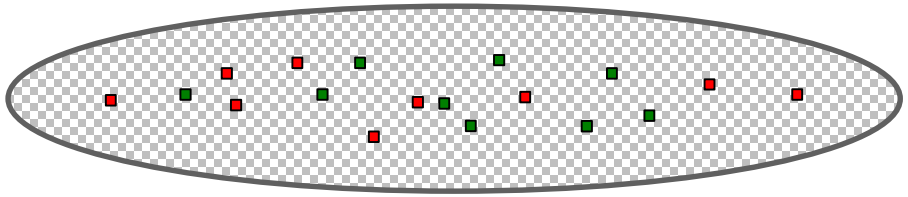
Step 1 Outputs

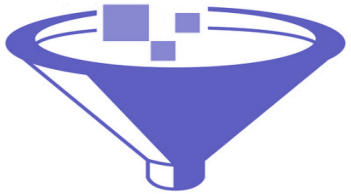
Step 2:
Data Analytic Techniques

Step 2 Outputs

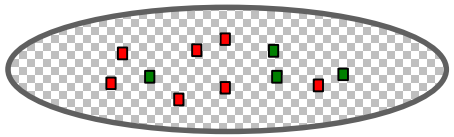
Step 3:
Prioritization

Prioritized Notable Items

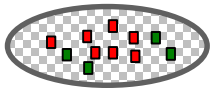


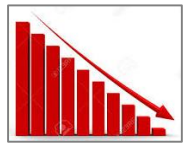


Apply a set of filters to examine significant risks (i.e., What Could Go Wrong)
(e.g., duplicate payment)



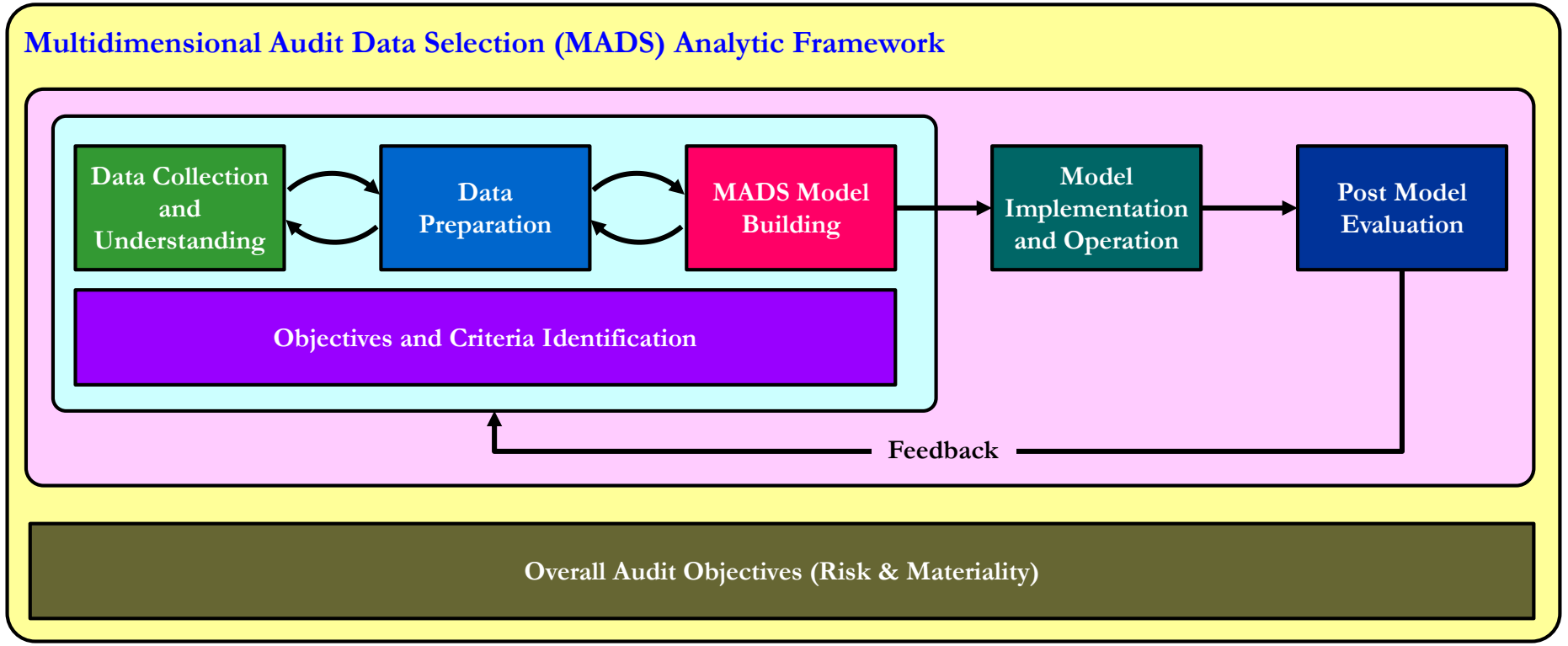
- Additional Filters
- Visualization Techniques
  (e.g., scatter plots)
- Professional Judgement
  (e.g., knowledge and experiences)
- Outlier Detection Techniques
  (e.g., classification & clustering).





- Use professional judgement based on the importance of each step 1 filter and step 2 filter.
- Use the step 1 and/or step 2 results.
- Use a reasonable factor
  (e.g., dollar amount).

MADS ANALYTIC FRAMEWORK

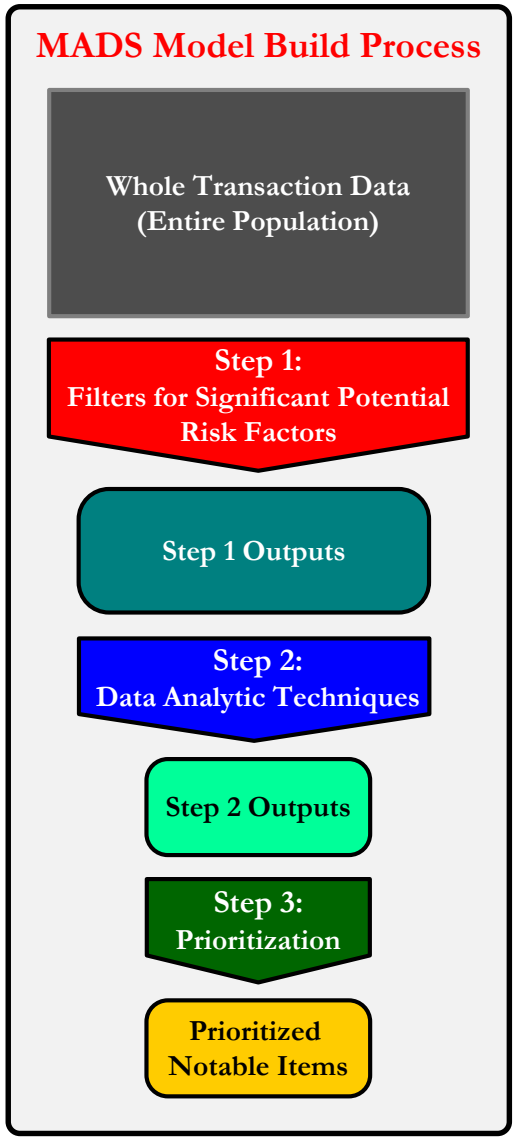


- Carry out empirical tests of whether or not the MADS process results in a more effective auditing process compared to the current sampling processes.
  - Revenue Cycle (Order-to-Cash)
  - Expenditure Cycle (Purchase-to-Pay)
  - Payroll Cycle
  - General Ledger

DATA DESCRIPTION

❖ Expenditure (Procure-to-Pay) cycle

- From Hub of Analytics Education (http://www.hubae.org)
- Bibitor LLC is a retail liquor chain company that sells wine and spirits.
  - 1 year dataset (6/21/2016 - 6/20/2017)
  - 2,291,725 records and 5,234 invoices



SUMMARY

RUTGERS
Rutgers Business School
Newark and New Brunswick

MADS Model Build Process

Whole Transaction Data
(Entire Population)

Step 1:
Filters for Significant Potential Risk Factors

Whole Transaction Data

2,291,725 purchase transactions records
5,234 invoices

ID	Potential Test (or Filters)
PUR-03	Identify purchases that are not properly approved (i.e., authorization limits - $250,000) by the authorizer (i.e., Chief Operating Officer).
PUR-06	Identify purchases that are received after payment.
PUR-09	Identify unusual purchases by producing exception reports of order amount/quantity that is too high (e.g., higher than percentile 95 value or greater than $5M/500 Units).
PUR-17	Identify purchases made to vendors who are not on the approved vendor list.
INV-02	Identify invoices where the order amount is different from the invoice amount.
INV-13	Identify multiple invoices at or just under approval cut-off levels (i.e., $250,000).
PAY-08	Identify payments that are made to invoices without purchase orders.
DUP-02	Identify duplicate invoices and/or amounts.

SUMMARY

RUTGERS
Rutgers Business School
Newark and New Brunswick

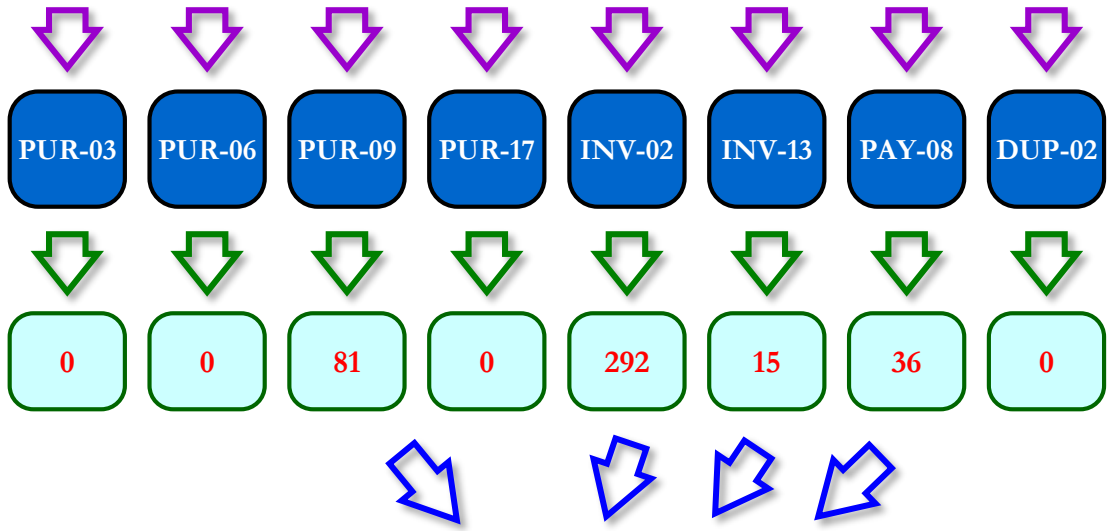
MADS Model Build Process

Whole Transaction Data
(Entire Population)

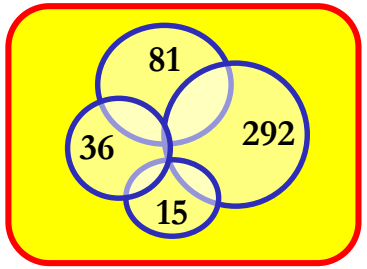
Step 1:
Filters for Significant Potential Risk Factors

Whole Transaction Data

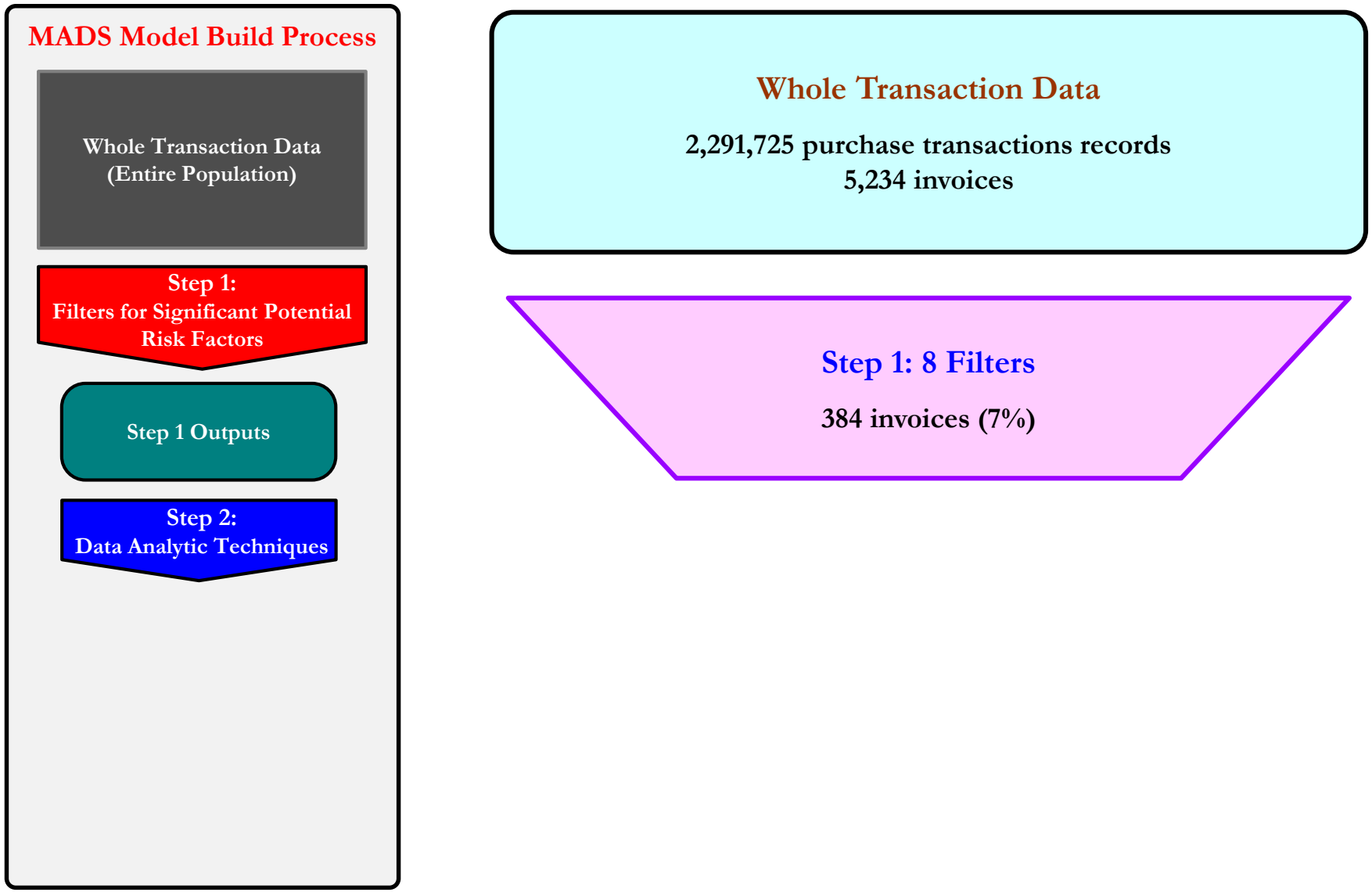
2,291,725 purchase transactions records
5,234 invoices



After applying step 1 filters, 384 invoices (7%) has been identified for further investigation.



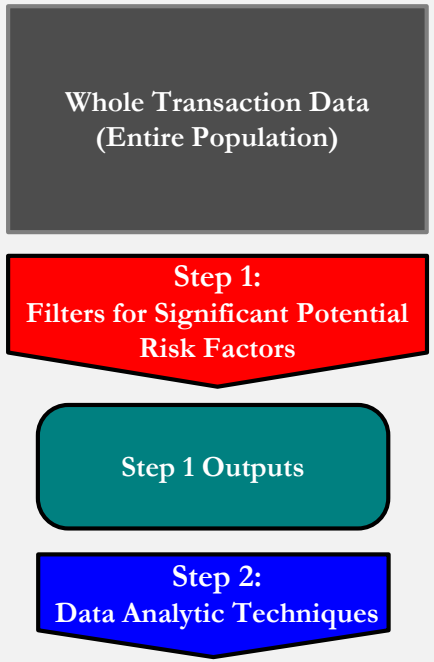
SUMMARY


MADS Model Build Process
Whole Transaction Data
(Entire Population)
Step 1:
Filters for Significant Potential
Risk Factors
Step 1 Outputs
Step 2:
Data Analytic Techniques
Whole Transaction Data
2,291,725 purchase transactions records
5,234 invoices
Step 1: 8 Filters
384 invoices (7%)

SUMMARY

RUTGERS
Rutgers Business School
Newark and New Brunswick

MADS Model Build Process



Whole Transaction Data

2,291,725 purchase transactions records
5,234 invoices

ID	What Could Go Wrong?	S_ID	Potential Test (or Filters)
PUR-09	Purchases are made with unusual order amount and quantity.	SUB-01	Actual price is larger than approved price.
		SUB-02	Invoice amount is significantly larger than order amount (> $100,000) *.
INV-02	Order amount does not match with invoice amount.	SUB-01	Actual price is larger than approved price.
		SUB-02	Invoice amount is significantly larger than order amount (> $100,000) *.
INV-13	Purchases are made just under approval cut-off amount.	SUB-01	Actual price is larger than approved price.
		SUB-03	Identify purchases made to vendors who are not in the approved vendor list.
PAY-08	Payments are made to invalid purchase orders.	SUB-04	Identify unusual payment without purchase orders (> $5,000) **.

* 1% of Performance Materiality
** Based on Judgement

SUMMARY

RUTGERS
Rutgers Business School
Newark and New Brunswick

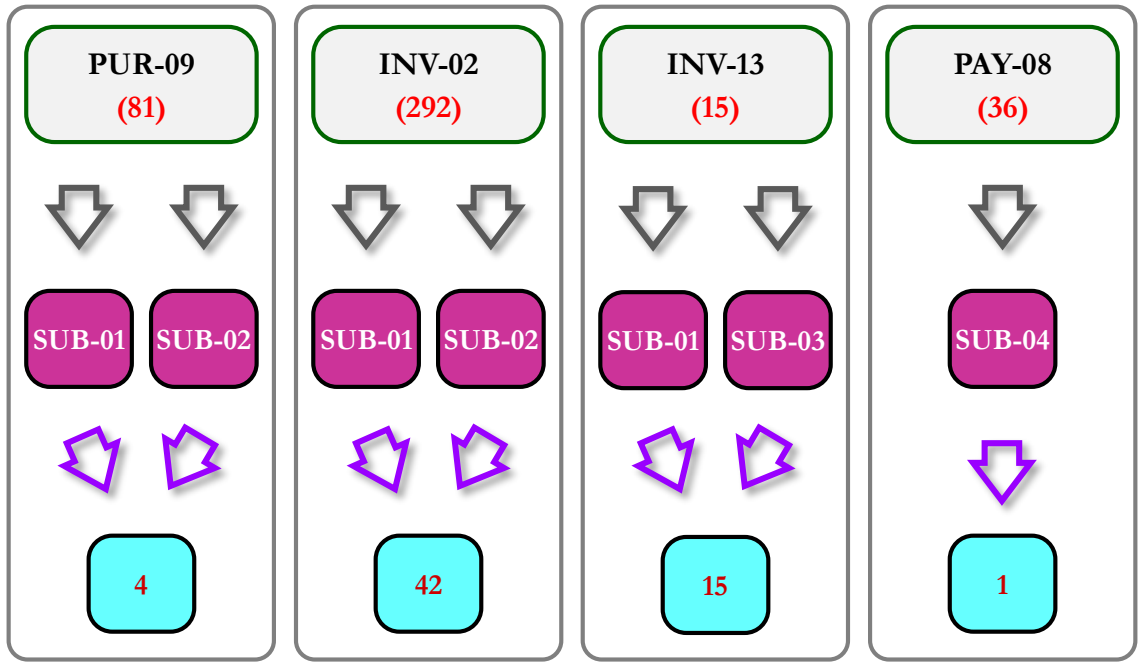
MADS Model Build Process

Whole Transaction Data
(Entire Population)

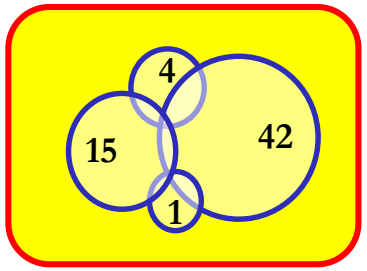
Step 1:
Filters for Significant Potential Risk Factors

Step 1 Outputs

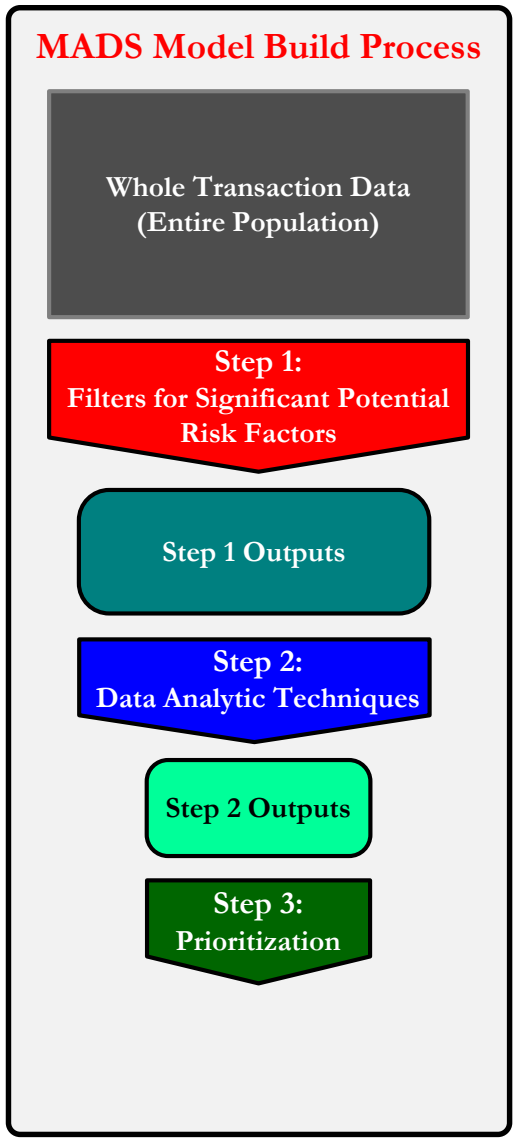
Step 2:
Data Analytic Techniques

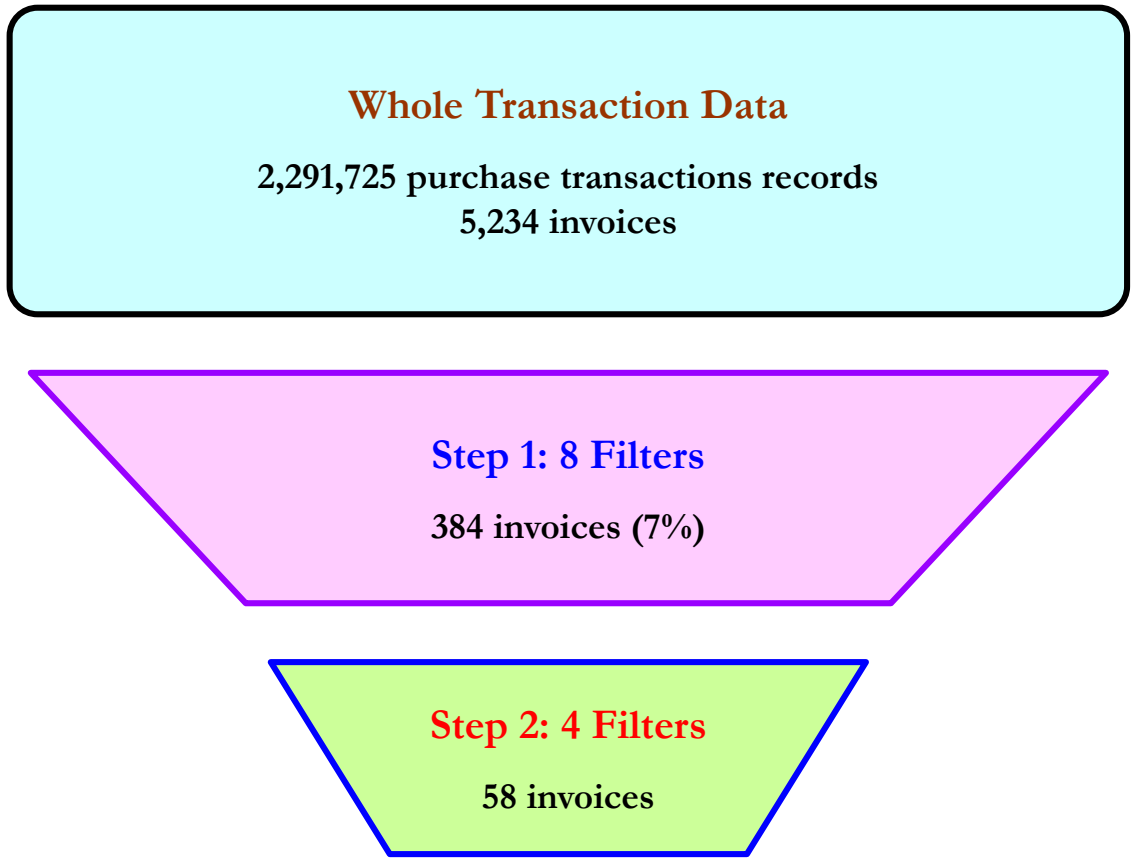


After applying step 2 filters, 58 invoices (out of 384 step 1 outputs) has been detected.

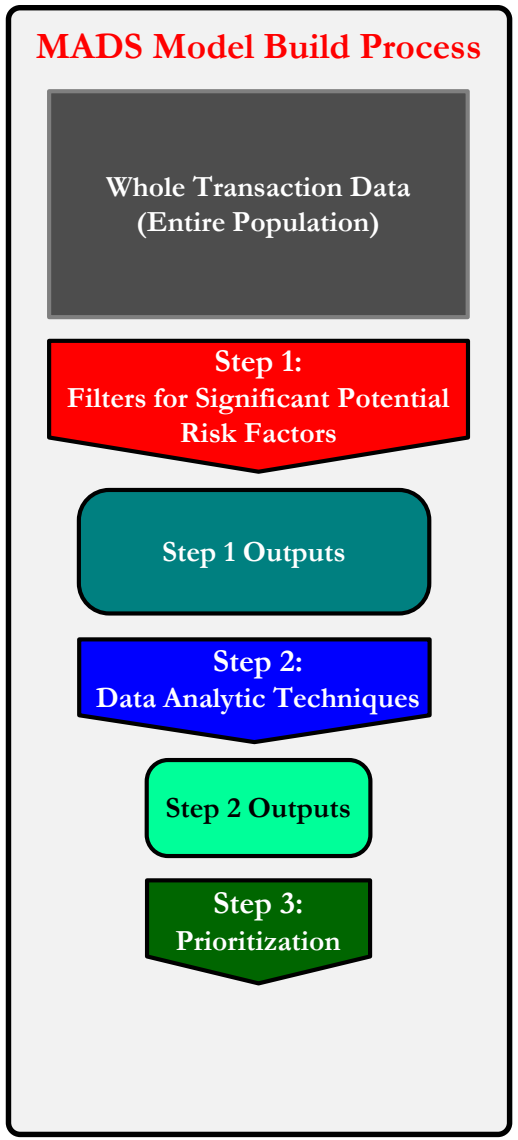


SUMMARY


MADS Model Build Process
Whole Transaction Data
(Entire Population)
Step 1:
Filters for Significant Potential
Risk Factors
Step 1 Outputs
Step 2:
Data Analytic Techniques
Step 2 Outputs
Step 3:
Prioritization


Whole Transaction Data
2,291,725 purchase transactions records
5,234 invoices
Step 1: 8 Filters
384 invoices (7%)
Step 2: 4 Filters
58 invoices

SUMMARY



Whole Transaction Data

2,291,725 purchase transactions records
5,234 invoices

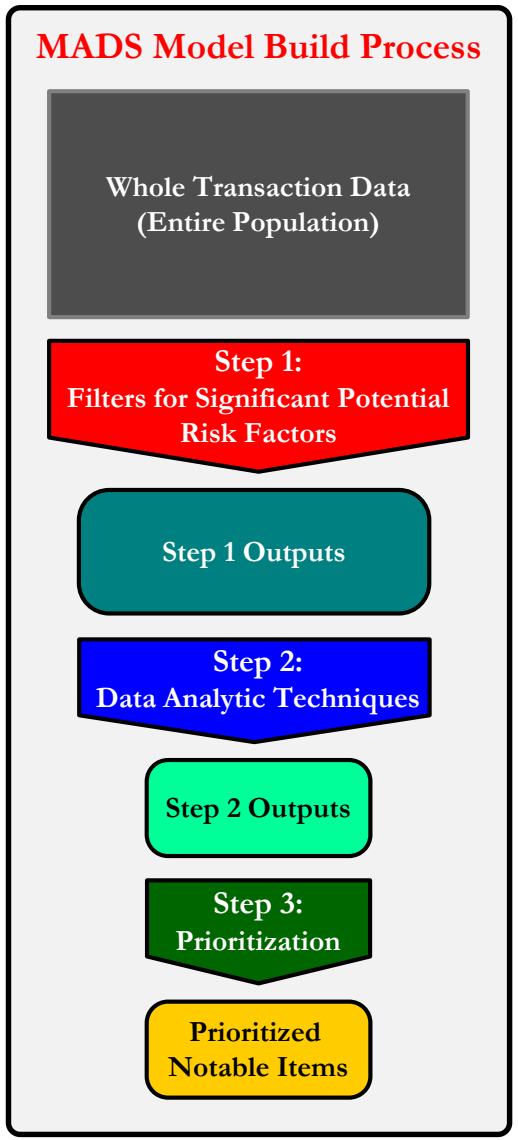
❖ Prioritization

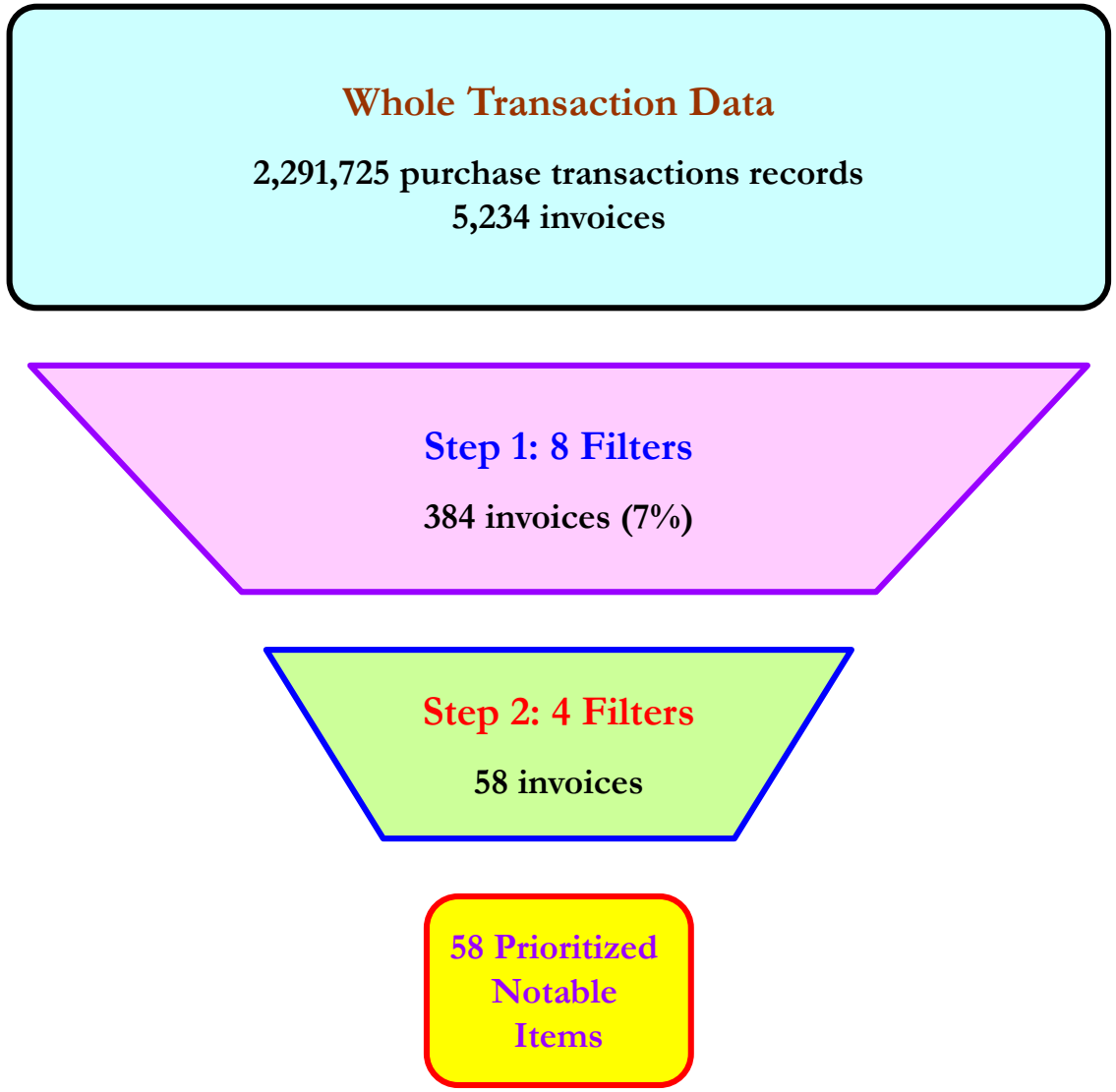
- Use the step 1 and step 2 results.
- Example
  - Assume that invoice #273 (one of 58 notable items) has three violations in step 1 and one violation in step 2, and the dollar amount is $ 265,000.
  - Violation score will be calculated as:

$$\frac{(\text{Step 1 Violations} + \text{Step 2 Violations})}{\text{Number of Step 1 and 2 Filter Applied}} = \frac{(3+1)}{(8+2)} = .4$$

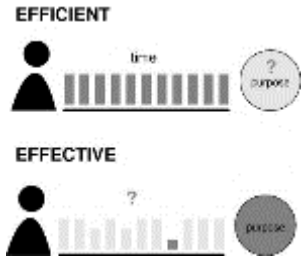
  - Suspicion score will be calculated as:
    Amount * Violation Score = 265,000 * .4 = 106,000.
- Using the suspicion score, step 2 outputs (i.e., 58 notable items) are prioritized.

SUMMARY

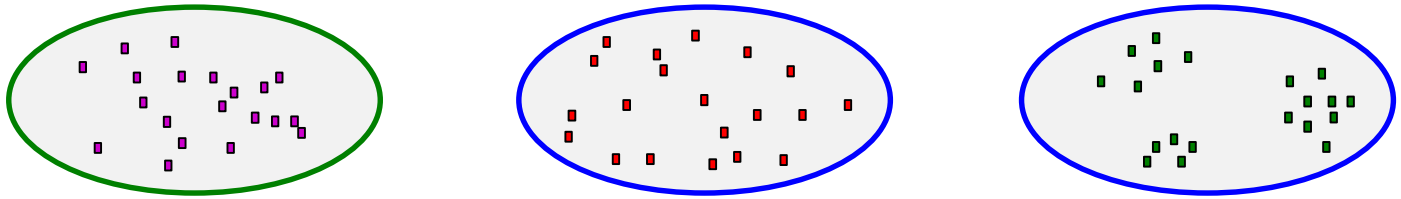

MADS Model Build Process
Whole Transaction Data
(Entire Population)
Step 1:
Filters for Significant Potential
Risk Factors
Step 1 Outputs
Step 2:
Data Analytic Techniques
Step 2 Outputs
Step 3:
Prioritization
Prioritized
Notable Items


Whole Transaction Data
2,291,725 purchase transactions records
5,234 invoices
Step 1: 8 Filters
384 invoices (7%)
Step 2: 4 Filters
58 invoices
58 Prioritized
Notable
Items

EVALUATIONS



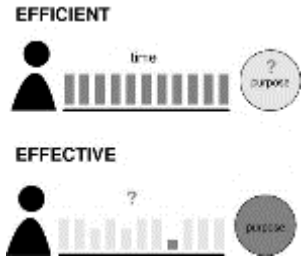
- Statistical Sampling vs. Non-statistical Sampling vs. MADS
- Two Aspects
  - Effectiveness - More suspicious items (i.e., errors) in the sample
  - Efficiency - Less sample size (?)
- Three Potential Evaluation Methods
  - Method 1: Benchmark (based on the assumption that we have already identified filters which can discover all errors in the full population)
  - Method 2: Random Transaction Changes
  - Method 3: Realistic Error Seedings by Experienced Auditors (Preferred)



Benchmark

MUS (Statistical) vs. Non-statistical vs. MADS

EVALUATIONS



- Statistical Sampling vs. Non-statistical Sampling vs. MADS
- Two Aspects
  - Effectiveness - More suspicious items (i.e., errors) in the sample
  - Efficiency - Less sample size (?)
- Three Potential Evaluation Methods
  - Method 1: Benchmark (based on the assumption that we have already identified filters which can discover all errors in the full population)
  - Method 2: Random Transaction Changes
  - Method 3: Realistic Error Seedings by Experienced Auditors (Preferred)

	Method 1 – Benchmark	Method 2 – Random Transaction Changes
Pro	• Use original data set (no manipulation)	• Provide error population and error items
Con	• Based on the assumption that identified filters can detect all errors • MADS is inherently more effective than traditional sampling approaches since MADS filters are a subset of identified filters.	• Change the values of original data at random. • Random value changes may not represent realistic errors.

EVALUATIONS
METHOD 1 – BENCHMARK

RUTGERS
Rutgers Business School
Newark and New Brunswick

❖ Benchmark

- Apply all 27 filters (8 step 1 filters + 19 additional filters).
- Assume that items filtered by 27 filters are all errors in the entire population.
- 539 (10.3%) items are identified and regarded as benchmark.

Population	$ 306,093,663
Tolerable Misstatement (75% of OM) (OM: 5% of Total Revenues)	$ 15,494,054
Expected Misstatement (1%)	$ 3,060,937
Risk of Incorrect Acceptance	10%
Sample Size	67

❖ Monetary Unit Sampling (MUS)

- Use CaseWare IDEA.
- Identify 67 items.

❖ Non-statistical Sampling

- Based on judgement, stratify items into four groups.
- All 23 large items (>= $ 1M) are included.
- 44 items are randomly selected from the arbitrarily allocated three groups (i.e., 50%, 30% and 20%).

Amount	Allocation	
>= $ 1M		23
>= $ 250,000	50%	22
>= $ 100,000	30%	13
>= 0	20%	9
Sample Size		67

❖ MADS

- All 58 notable items are selected.

EVALUATIONS

METHOD 1 – BENCHMARK

RUTGERS
Rutgers Business School
Newark and New Brunswick

Sampling Risk (10%)	MUS	Non-statistical	MADS	Benchmark
Sample Size	67	67	58	5,234
No. of Error Items	13 (2.4%)	23 (4.3%)	58 (10.7%)	539 (100%)
No. of Violations Detected	20 (2.6%)	35 (4.6%)	125 (16.6%)	751 (100%)

☑ More Effective

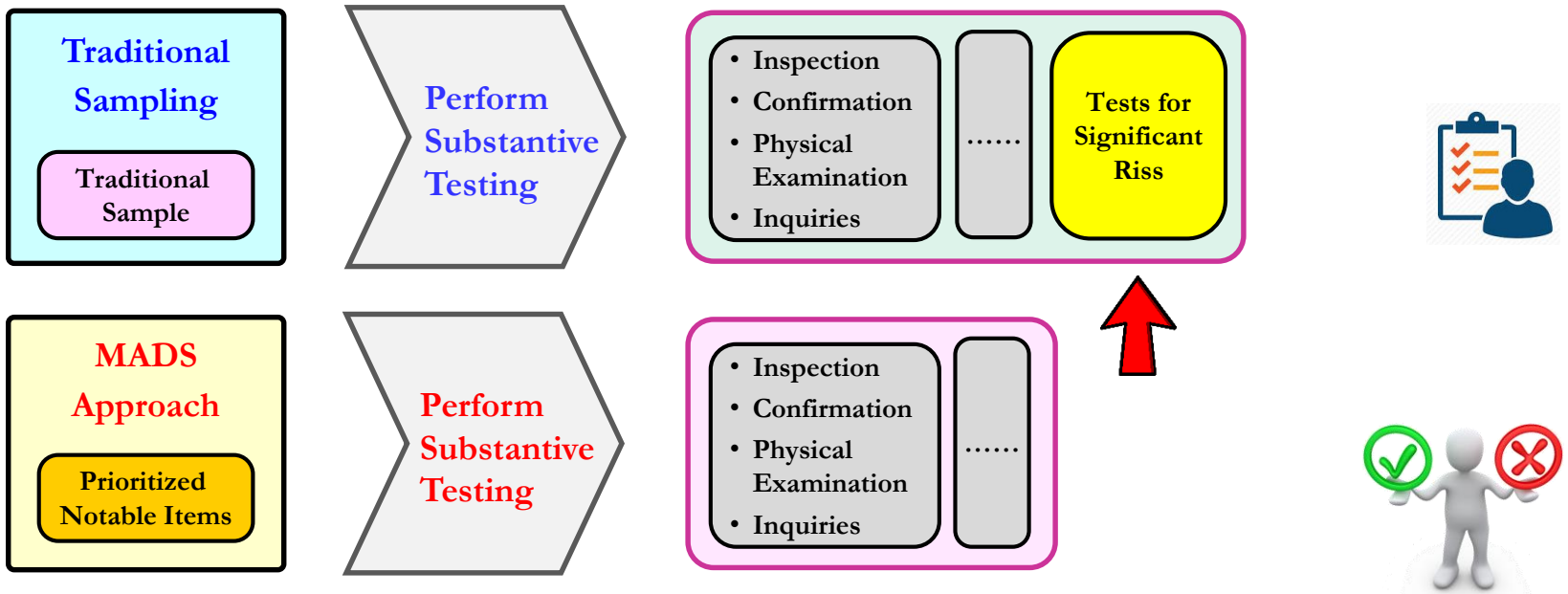
☑ More Efficient

EVALUATIONS

METHOD 1 – BENCHMARK

RUTGERS
Rutgers Business School
Newark and New Brunswick

Sampling Risk (10%)	MUS	Non-statistical	MADS	Benchmark
Sample Size	67	67	58	5,234
No. of Error Items	13 (2.4%)	23 (4.3%)	58 (10.7%)	539 (100%)
No. of Violations Detected	20 (2.6%)	35 (4.6%)	125 (16.6%)	751 (100%)



EVALUATIONS

METHOD 2 – RANDOM TRANSACTION CHANGES

- Values are randomly changed at the purchase transaction level.
  - Total number of value changes
    - 0.01% of total purchase transactions (2,291,725): 225 transactions
  - The values of amount, price, receiving date, and pay date are randomly changed.

Whole Transaction Data

2,291,725 purchase transactions records
5,234 invoices

Step 1: 8 Filters

481 invoices (9.2%)

Step 2: 4 Filters

99 invoices

99 Prioritized Notable Items

EVALUATIONS

METHOD 2 – RANDOM ERROR SEEDINGS

Sampling Risk (10%)	MUS	Non-statistical	MADS *
Sample Size	67	67	67
No. of Suspicious Items	4 (3%)	5 (3.8%)	24 (18%)
No. of Errors	133 Items		

* For comparison, choose top 67 items from 99 notable items.

☑ More Effective

☑ More Efficient

NEXT STEPS

- Different Data Analytics Techniques in Step 2 (e.g., Clustering)
- Different Prioritization Criteria (e.g., Professional Judgement)
- Different Evaluation Methods (e.g., Manual Realistic Error Seedings)
- Different Data Sets




QUESTIONS & SUGGESTIONS