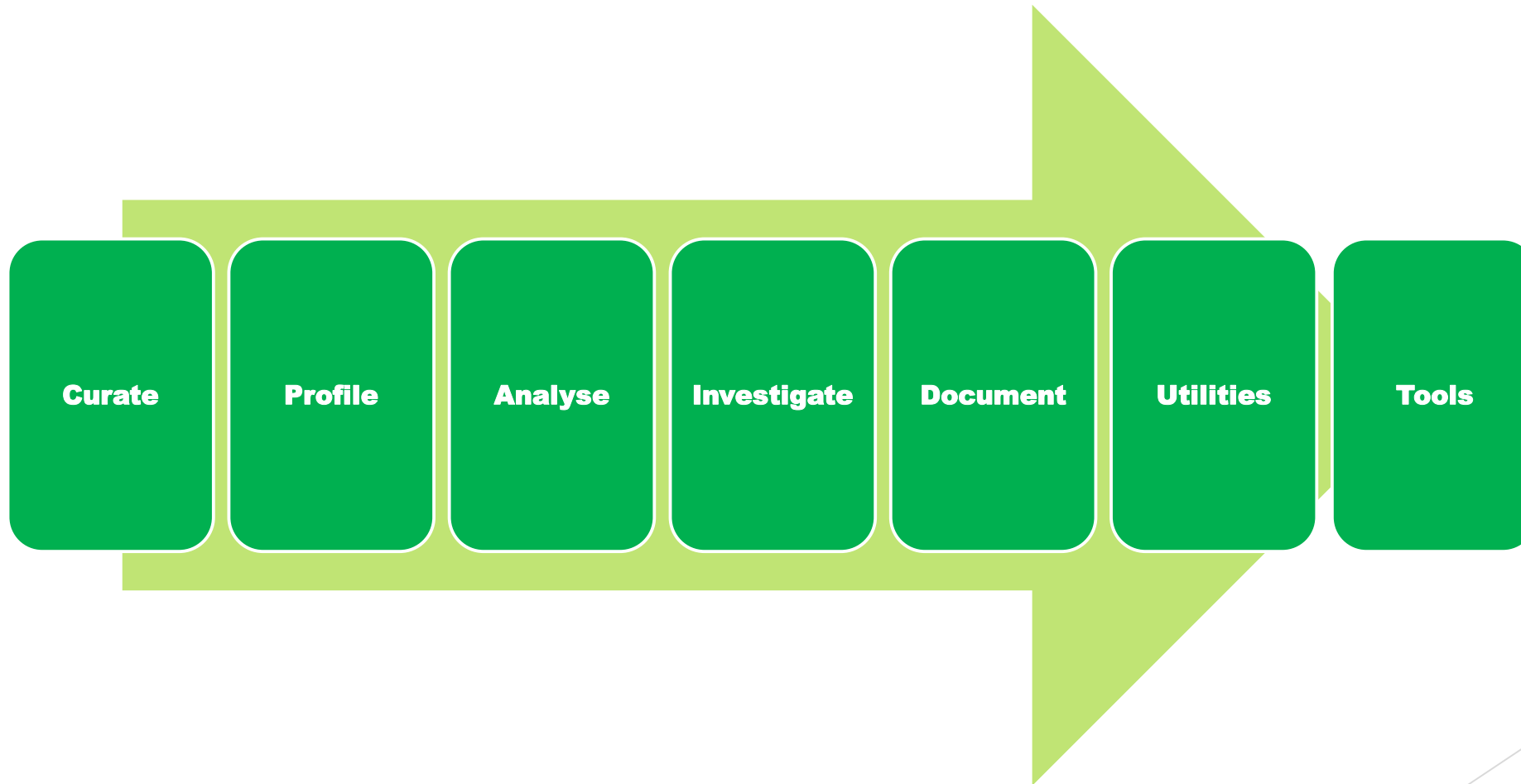# Demo of
Key functions of Data Analytics using eCAAT

CA. Abdul Rafeq, FCA, CISA, CGEIT, CIA

Managing Director, Wincer Infotech Limited

www.wincaat.com

Most Common Features of Data Analytics

Curate | Profile | Analyse | Investigate | Document | Utilities | Tools

Column Statistics

Details: Column Statistics

Displays Column wise Statistics of numeric, date and character data.

Example:

Review Column Statistics of selected file to match with source data.

Use File:

sales_trans or any file from eCAAT's sample data

Assertions:

Cut off / Completeness

Stratify: Numeric

Stratify: Numeric

Groups numeric data into different Strata based on specified intervals and displays count and sum with % for each strata.

Example:

Stratify Sales by amount to identify the sum and count as per each strata.

Use File:

Salestrans file from eCAAT's Sample Data

Assertions/Analysis:

Accuracy / Completeness

Summarise: Quick

Summarize: Quick

Summarises/totals the data in the selected numeric column based for each distinct value and displays corresponding count and sum with %.

Example:

Summarise sales as per unique value to identify pattern of sales based on value.

Use File:

Salestrans file from eCAAT's Sample Data

Assertions/Analysis:

Accuracy / Completeness

Profile: Percentiles

Profile: Percentiles

Groups values for the selected column as per clusters based on no. of clusters specified.

Example:

Identify purchases as per clusters based on quantity of purchase.

Use Files:

Procurement file from eCAAT Sample Data

Assertions/Analysis:

Accuracy / Completeness

Duplicates in Same Row

Duplicates In Same Row

Displays Duplicate Values (Character, Date or Numeric) in the Same Row.

Example:

Find transactions entered and authorised by same user.

Use Files:

procurement file from eCAAT's sample data

Assertions:

Authorisation / SOD

Character Gaps

Identify Gaps: Character

Displays gaps in Successive Alpha-Numeric records whose difference is within the Given Character Mask.

Example:

Identify Gaps in Transaction No. which is alpha-numeric field ignoring the character.

Use File:

AccountTrans file from eCAAT's sample data

Assertions:

Accuracy / Validity

Same Same Different

Same Same Different

Matches values of each column of all rows to identify records where the values from two selected columns are same but value of third column is different.

Example:

Identify suppliers whose Name & PAN are same but addresses are different.

Use File:

procurement file from eCAAT's Sample data

Assertions:

Accuracy / Validity

Outliers

Identify Outliers: Numbers

Displays records that exceeds 'x' no. of times, average or std. deviation of selected numeric column.

Example:

Identify sales which are 3 times average sales/std. deviation.

Use File:

SalesTrans File from eCAAT's sample data

Assertions/Analysis:

Accuracy / Exceptions

Sampling: Stratified Random

Sampling: Random – Profile - Stratified on Date

Perform date stratification to Select specific number of records on random basis within each specified Date band after date Stratification.

Example:

Identify random no. of samples as specified from each strata of date-related data.

Use File:

SalesTrans file from eCAAT's Sample Data

Assertions:

Authorisation / Validity /SOD

Fuzzy Match: Arrange

Fuzzy Match: Arrange

Displays similarities in two rows based on values by eliminating transposition within cell values with option to identify duplicates after arranging data.

Example:

Find duplicate transactions which have been entered in different ways by changing order of words.

Use File:

AccountTrans file in eCAAT's Sample data

Assertions:

Accuracy / Completeness

Fuzzy Match: Normalise

Fuzzy Match: Normalize

Removes Special Characters, Multiple Spaces, Leading and Trailing Spaces in a Column.

Example:

Identify similar narrations by removing special characters.

Use File:

AccountTrans file from eCAAT's Sample Data

Assertions:

Accuracy / Validity

Fuzzy Match: Similar

Fuzzy Match: Similar

Compares two values and applies transformation to find values/data that are similar.

Example:

Identify values which are similar for selected column after applying the specific difference between columns

Use File:

Dr&CrColumn file from eCAAT's Sample Data

Assertions:

Accuracy / Validity

Fuzzy Match: Soundex/Soundslike

Fuzzy Match: Soundex/SoundsLike

Soundex displays the four digit Soundex code for selected for columns to identify exact/approximate phonetic similarities.
SoundsLike displays similarities of words from two columns based on exact/approximate Soundex codes.

Example:

Use the Soundex/Soundslike function to identify similar sounding names entered with different spellings.

Use File:

Soundex/Soundslike file from the eCAAT's Sample Data

Assertions:

Accuracy / Completeness

Analysis: Cluster

Analysis: Cluster

Displays summary of groups of data from selected set of 2 columns such that each cluster has data of similar characteristics as specified.

Example:

Identify specific group of records for each unique material code based on quantity and unit price as per specified clusters.

Use File:

procurement file from eCAAT Sample Data.

Assertions:

Validity / Completeness

Investigate: MIS

Investigate: MIS

Displays for selected column: Count, Total, Minimum, Maximum, and Average of Records for the Selected Period.
Multiple Worksheets can be generated at once for various options.

Example:

Generate dynamic 'MIS' for Sales based on day, day of week, monthly, quarterly sales.

Use File:

SalesTrans File from eCAAT's sample data.

Assertions:

Accuracy / Completeness

Join Files

Relate: Join Files

Displays records after joining data from two worksheets based on specific common/uncommon column values.

Example:

Join 2 related files with unique identifier to identify sales beyond authorization limit.

Use Files:

Sales Trans & Sales_Auth Limit Files from eCAAT's Sample Data

Assertions:

Accuracy / Authorisation

3-Way Matching

Relate: 3-Way Matching

Displays records after joining data from upto 3 worksheets based on common/uncommon column values.

Example:

Identify specific fields from 3 related files on various parameters such as availability in all, selected files.

Use fFiles:

Exp_Apr, Exp_May and Exp_Jun from eCAAT's Sample Data.

Assertions:

Accuracy / Authorisation

Authentication Check

Templates: Authentication Check

Displays the records from two Worksheets that match two conditions.

Example:

Identify sales entered beyond authorization limit by each of the users.

Use Files:

SalesTrans and AuthLimit from eCAAT's Sample Data.

Assertions:

Authorisation / SOD

Compare Files

Compare Files

Displays the count, sum or difference of two or more similar columns from two different worksheets.

Example:

Compare Sales between two given periods based on sum and count for each customer.
Use files:

Invoice 2016 and Invoice 2017 from eCAAT's sample data

Assertions:

Accuracy / Completeness

Analytical Review

Templates: Analytical Review

Displays difference between values of two numeric columns in number and in percentage.

Example:

Compare Balance Sheet items for two years and identify variance between them in value & %.

Use File:

BalanceSheet file from eCAAT's sample data

Assertions/Analytical Procedure:

Accuracy / Presentation & Disclosure

Append Files: Matched Columns

Relate: Append Files

Displays records after Appending multiple worksheets with same record structure. (matching columns

Example:

Append expenses of 3 months in a single worksheets.

Use Files:

ExpApril, ExpMay & ExpJune files from eCAAT's Sample Data

Assertions:

Accuracy / Completeness

Relate: Consolidate

Relate: Consolidate

Merges data from selected worksheets on specific column and displays count and sum.

Example:

Consolidate expenses of 3 months in a single worksheets.

Use Files:

ExpApril, ExpMay & ExpJune files from the eCAAT's Sample Data

Assertions:

Accuracy / Completeness

Relative Size Factor

Relative Size Factor (RSF)

Displays the variation between highest value and second highest value in terms of difference and multiple for specified numeric column.

Example:

Identify differential price and ratio between first maximum and second minimum price for each material code.

Use File:

procurement file from eCAAT's sample data

Assertions:

Correctness / Validity

Maximum Variance Size Factor

Maximum Variance Factor (MVF)

Displays the variation between highest value and lowest value in terms of difference and multiple for specified numeric column.

Example:

Identify differential price and ratio between highest and lowest price for each material code.

Use File:

procurement file from eCAAT's sample data

Assertions:

Correctness / Validity

Analysis: Quadrant

Can you help the business manage at speed?

Questions?

rafeq@wincaat.com

Thank you!