

# ACEM MICROCRÉDIT MONTRÉAL

McGill MMA Student Consulting Project

Presented By: Leah XinRui Ma, Jungkang Zhang

Professor: Shoeb Hosain

# Agenda INTRODUCTION PAIN POINTS & USE CASE **USER PERSONA DATA PREPARATION & ANALYTICS SOCIAL IMPACT VALUE METRICS SOLUTION ARCHITECTURE DASHBOARD TECHNICAL EVOLUTION**

**CONCLUSION & APPENDIX** 

# Introduction

## About ACEM - Business Overview



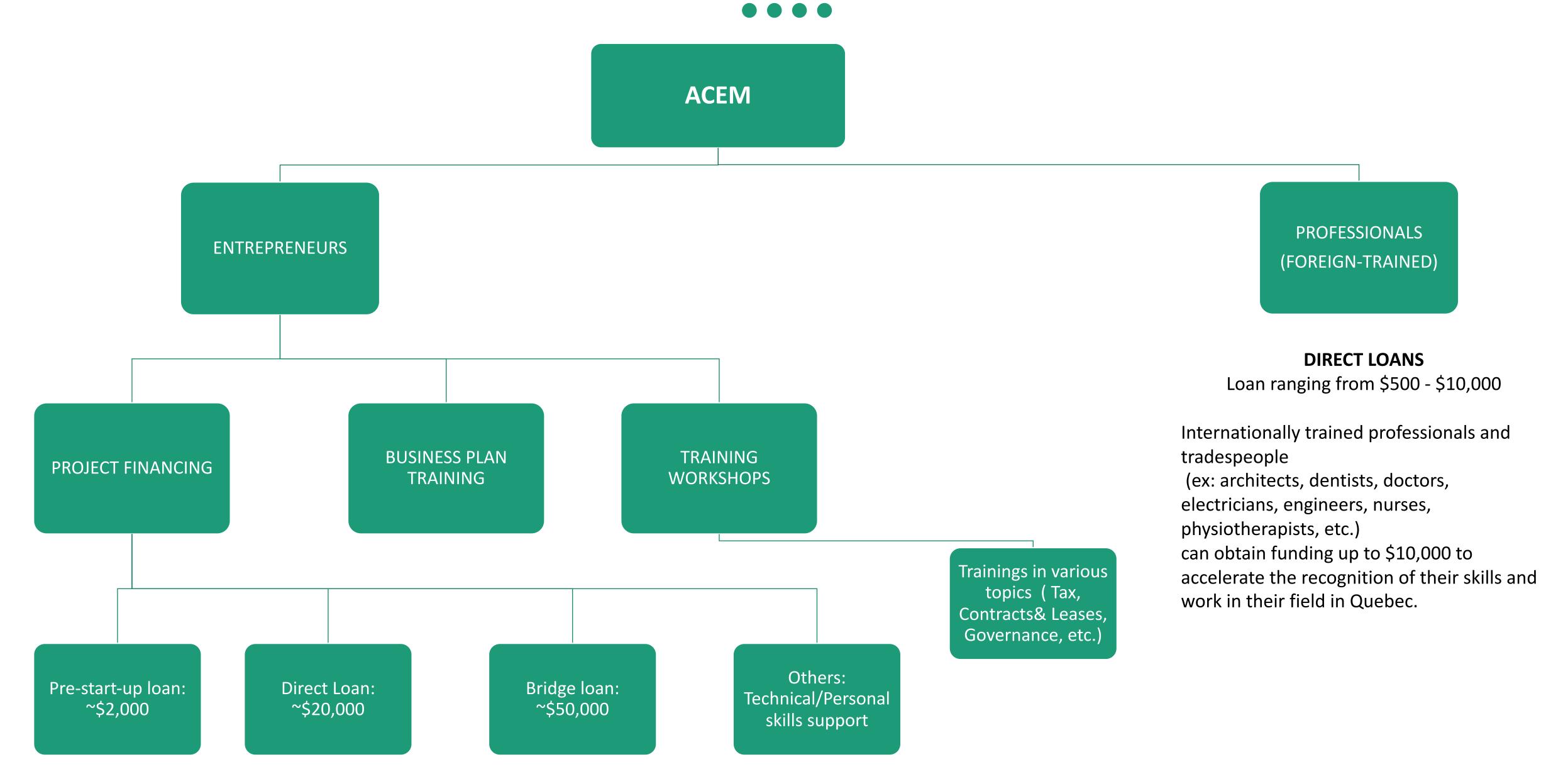
Since 1990

ACEM Microcrédit Montréal is a social enterprise who seeks to democratize financing through the participation of a network of investors and borrowers to promote an inclusive and prosperous society.

ACEM serves low-income groups, predominantly women, unemployed youth, immigrants, and visible minorities.



## Loan Programs Structure



# Use Case



## Project 1 - Pain Points

#### **Description**

For ACEM project 1, in order to attract and generate more donations from board members, funders, partners, donors, etc., it is necessary to represent the work ACEM do into societal impact and metrics for which the investors care.

- Results Compiled Manually
  - Analysis and reports often compiled manually
- Difficult to Quantify Societal Impact
  Societal impacts are difficult to be represented quantitatively
- Lacking Dynamic Graphical Representation of the Metrics
  Insufficient visual information to help donors to understand how their investments and donations are contributing to the community



## Project 1 – Use Case

#### **Description**

For ACEM project 1, in order to attract and generate more donations from board members, funders, partners, donors, etc., it is necessary to represent the work ACEM do into societal impact and metrics for which the investors care.

- R & Python Scripts for generating Analysis Templates

  Hash information to protect privacy & scripts for Excel templates
- Suggested Value Metrics for Societal Impact

  Metrics using existing data and suggestive future data
- Using MicroStrategy Desktop. Aim to minimize cost and also act as initial starting point.

## User Persona

## The Donors / Investors



Individuals (41%)

Most of the investors are Individuals

→ Metrics needs to be easily understandable



**Commercial Businesses (33%)** 

Such as Desjardins and other businesses

→ Metrics needs to be showing impact on people whom may be their future potential customers



Other Ethic Groups (26%)

Such as Religious Institution, Ethical Investors, NGOs

→ Want the process to be transparent and know about detail impact on minority groups of concern

<sup>\*</sup> Appendix 1 for detailed breakdown of donors

# Data Preparation



## Data Preparation Overview

Portfolio
Analyses,
Reports, etc 6 Files

Structured

Data - 15 Files





#### Loan progress tracking

Structured data frames were merged and used to calculate conversion rate at each stage of the Entrepreneur Program



#### **Predictive Modelling**

The team utilized these data files to construct a predictive model which can estimate number of jobs an entrepreneur can create in the future.



#### **New KPIs**

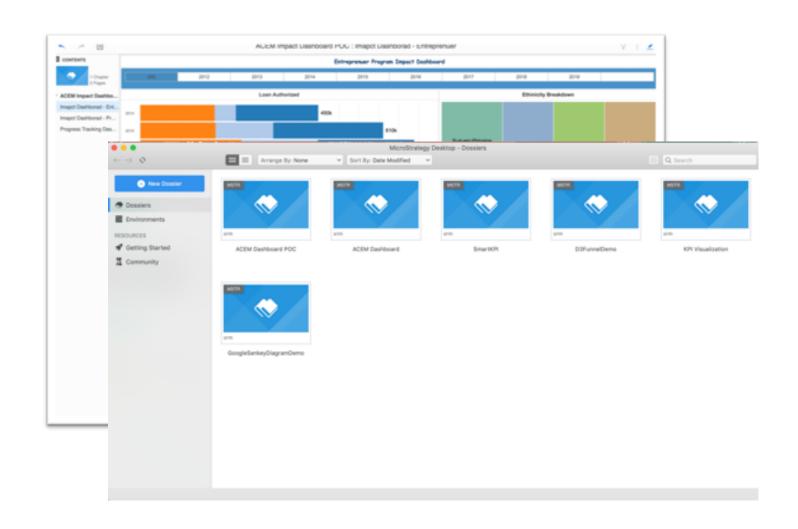


#### **Auto Data Reformatting**

Use Python to automatically format data frames for the dashboard use



#### **Import into Dashboard**



## Data Pre-processing





#### **Data Anonymization**

Designed and implemented an automated hashing function in Python to automatically convert client privacy information into unique identifiers



#### **Data Translation**

Utilize Microsoft built-in functions to translate files in French into English



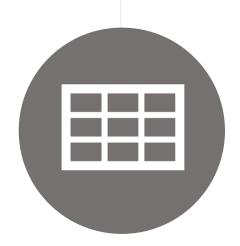
#### **Data Cleaning**

Remove invalid data entries, fill

NA values and cut unrelated

columns to validate the data for

analytics



#### Consolidation

Merge data frames at various
stages to further analyze and
construct ACEM Impact Dashboard
POC (to be presented)

## Data Pre-processing: Anonymization

Guarantee the same ID for a person across different files!

∠apo Hyacinthe	Aka
Yurani	Velez
Ysaac	Pango Rondon
Ysaac	Pango Rondon
Ysaac	Pango Rondon



## Advantage

- Ensure client's privacy is secured during outsourcing
- Automated code guarantees efficiency and accuracy



Zabo Hyacinthe	Aka	dd187f925ced277ebb4a6c03aa090269677fe117
Yurani	Velez	60f6534ebdf6ebffe806f1c7ce3644360374c4a2
Ysaac	Pango Rondon	354c43d9d66599e55e4be1365b8253d6ff8c99e5
Ysaac	Pango Rondon	354c43d9d66599e55e4be1365b8253d6ff8c99e5
Ysaac	Pango Rondon	354c43d9d66599e55e4be1365b8253d6ff8c99e5

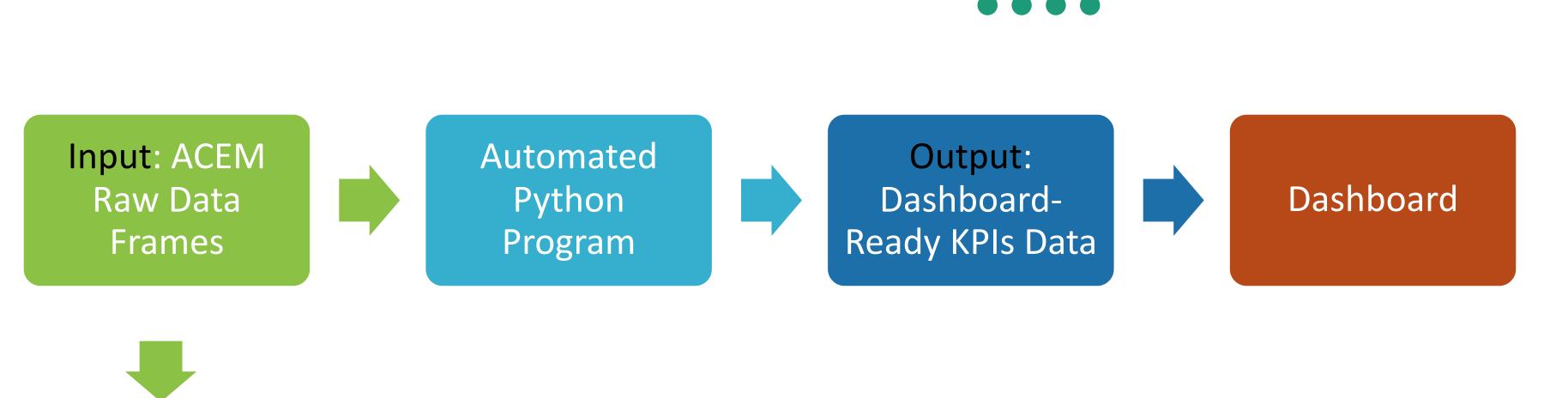


Delivery

 The hashing python file will be provided to ACEM with a user manual for future uses

# Data Analytics





1.Multiple Original Excel Files are imported as inputs
2. They are validated and utilized to generate KPIs in the Python program.

Data Frames	Size	Description
Service by program OE 2012- june 2019 OE.xlsx	(1198, 18)	Information about individuals who received any service
Analysis of business plan no loan OE.xlsx	(617, 8)	Entrepreneurs who didn't receive a loan at the end
Individual follow up OE.xlsx	(165, 8)	ACEM's follow up info after loan lent or no loan lent
Information session OE.xls	(1114, 8)	Entrepreneurs who attended the info sessions held by ACEM
Loans by program.xls	(191, 25)	Those who have received loans in this program
Training received by entrepreneurs OE.xls	(175, 8)	Records on those who have participated ACEM's trainings



Input: ACEM
Raw Data
Frames



Automated
Python
Program



Output:
DashboardReady KPIs Data



Dashboard

\*Please refer Appendix 2 to see the instructions of this Python tool

\*Please refer "df\_automation.py" file if you are interested in coding and the algorithms behind this tool (Within Junkang's Submission)

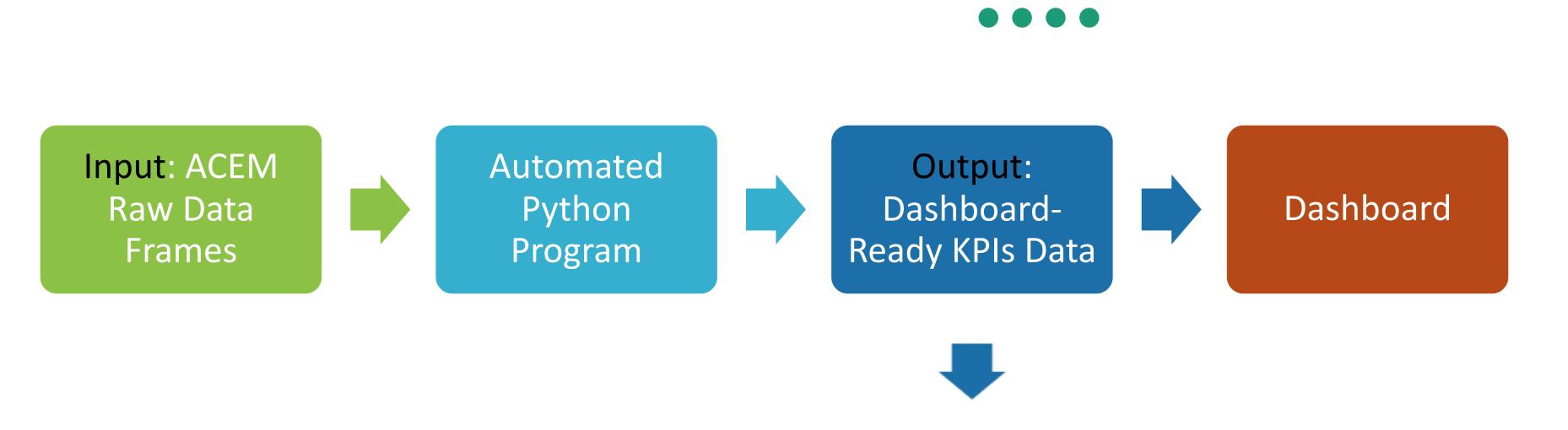


#### Why this step? & Its advantages

The aim of this step is to bridge the needs of cleaning and reformatting the raw data, transforming it into useful KPIs, and extracting key features in an efficient and automated way.

#### **Advantages:**

- Python is open-sourced and cost free.
- Python is utilized to automate the process in replace of repetitive labor work, resulting in 90% time saving for ACEM staff.
- The Python tool is set up in a fully-structured way, new requirements can be easily integrated into it without spending significant time.
- The technology stack further extends to the MicroStrategy dashboard, which means the dashboard auto-detects Python outputs and updates its content accordingly.



Four csv files are automatically exported from Python and each of them forms a function in the dashboard

\*Please refer Appendix 4 to see how the automation process works between Python and

MicroStrategy

Data Frames	Description
Entrepreneur Program.csv	Information about entrepreneurs who received any service
Professional Program.csv	Information about individual professionals who did/n't receive loan
Progress Tracking.csv	ACEM's follow up info after loan lent or no loan lent
Loan Portfolio by Year.xlsx	Entrepreneurs and professionals loan analysis info summarized by year from 2014 to 2019

<sup>\*</sup>Please refer Dashboard-read data.zip for data frame above (Within Junkang's Submission)







Automated Python Program



Output:
DashboardReady KPIs Data



Dashboard

\*Please refer to Appendix 3 for MicroStrategy Desktop Installation

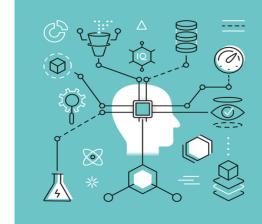


#### **Dashboarding Tools: MicroStrategy Desktop**

- By importing the four dashboard-ready KPIs dataframes into the ACEM Social Impact Dashboard template provided, all KPIs and metrics will be automatically generated and visualized
- After linking the dataframes by year, it further enables dynamic filtering across charts and graphs, providing interactivity and flexibility for interpreting insights and results



## Data Analytics: Machine Learning to ACEM Data





#### What is this model?

- It is a machine learning model built for ACEM's entrepreneur program, which predicts how many jobs an entrepreneur will be able to create in future based on current information.
- This model also lists the most influential factors affecting jobs created.

#### Why this model?

This model can be a good supplementary for ACEM when considering an applicant's qualifications and validity (Number of Jobs created matters the most for a community start-up)

#### Note:

Due to limited information given, this model is only a PoC and does not guarantee accurate prediction every time.

! It can be improved with more information and individuals records collected

Model: Number of Jobs Created by an Entrepreneur Prediction

#### Part 1: Predictions

Actual Values	Predictions	Income	Hours	Loan Amount
2	12.85	16000	122.2	15500
3	1.96	0	278.69	10000
1	1.41	18000	58.58	2000
2	7.15	7000	0.25	2000
0	1.01	7000	36.49	3400
2	1.38	35000	123.72	5000
1	1.27	7200	113.71	3365.18
1	10.28	7200	75.43	500
0	0.92	1500	7.5	2000
1	1.43	8400	489.3	10000
1	1.05	17000	166.75	1500
29	13.68	25000	188.48	13500

#### **Part 2: Important Features**

Loan Amount	0.420747
Cost Covered	0.186933
Tourisme et loisirs	0.152201
t_spent_fin	0.0820717
Hours	0.0690114
Income	0.0351377
Homme	0.0150536
Artistique et culturel	0.0110145
Restauration et secteur	0.0047491
Travailleur autonome	0.00416399
Anglais	0.00356266

\*Please refer Appendix 5 to see more details

# Social Impact Value Metrics

## Value Metrics 1: Higher Level Value Representations

#### **Easily Understandable Metrics for Individual Investors**



#### **Annual Loans Authorized**

Overview of the approximate amount



#### **Number of Entrepreneurs Supported**

Illustrating consistent effort in the field



#### **Average Age of Clients**

Demographic info:

Often in middle age where career development are becoming difficult



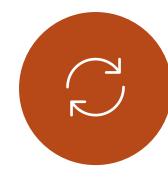
#### **Number of Professionals Supported**

Illustrating consistent effort in the field



#### **Education Level**

Demographic info to show the audiences that ACEM is serving



#### **Hours Spent Helping Clients**

Illustrating consistent effort in the field



#### **Family Status**

Highlighting
% of couples with children,
% divorced, etc.



#### % of Total Cost Covered

Extend of project being assisted by ACEM

### Value Metrics 2: Creating Jobs & Business Values

#### Impact on penitential customer for Businesses



#### **Jobs Created and Maintained**

People that successfully stepped into the workforce



#### Average years of working experience

Understanding of their career status so far



#### **Breakdown by Profession**

Impact are spread into various industries



#### **Employment Status**

Understanding of their career status so far

#### Value Metrics 3: Services to Minorities

#### Impact on minority groups for Religious Intuitions, NGOs & Ethical Investors



#### **Ethnicity Breakdown**

Highlighting minority group of concern



#### % of Low Income Minorities

Current income level that are lower than \$20,000\*



#### % of Woman

Empowering woman in business



#### **Progress Tracking**

Sample progress tracking for

Entrepreneur program

( can be expanded to other program)

Future
Suggestive Metrics
to
Collect & Consider





% Clients have stable income



% Clients have access to medical care



% Clients able to support their family



% Clients in the process / own a home





**Average Annual Profit** 



↑ Average number of employee per client business



% wages paid above poverty line





% Household with female as main income source



% Business run by woman



% Woman employed by businesses



% Jobs created by Woman Entrepreneurs

 $\bullet$   $\bullet$   $\bullet$ 





% Successfully start their business



**% Expanded their Business** 



% Invest in Education / Skills Enhancement

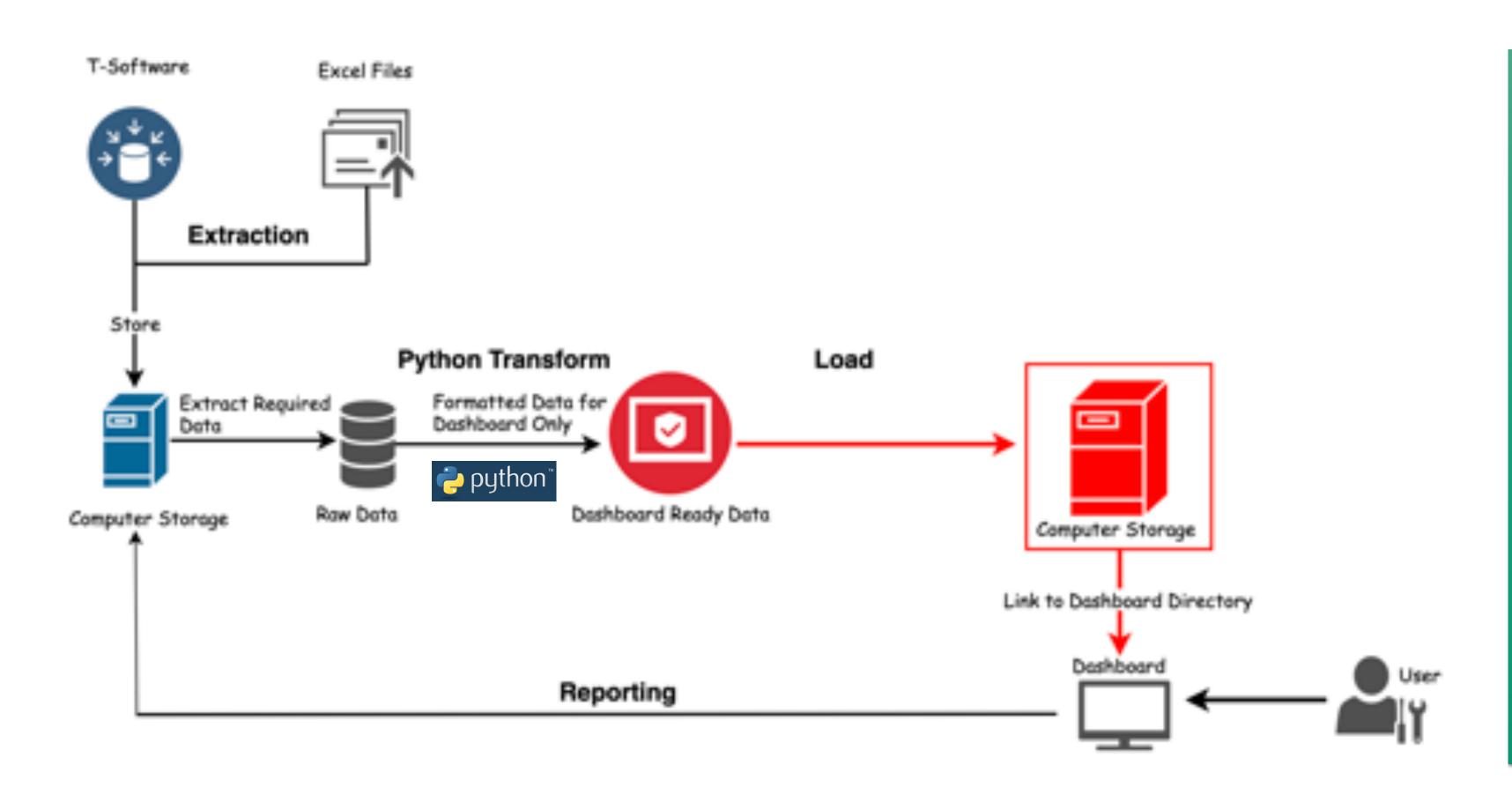


% Pay down debts

# Solution Architect

### Solution Architecture



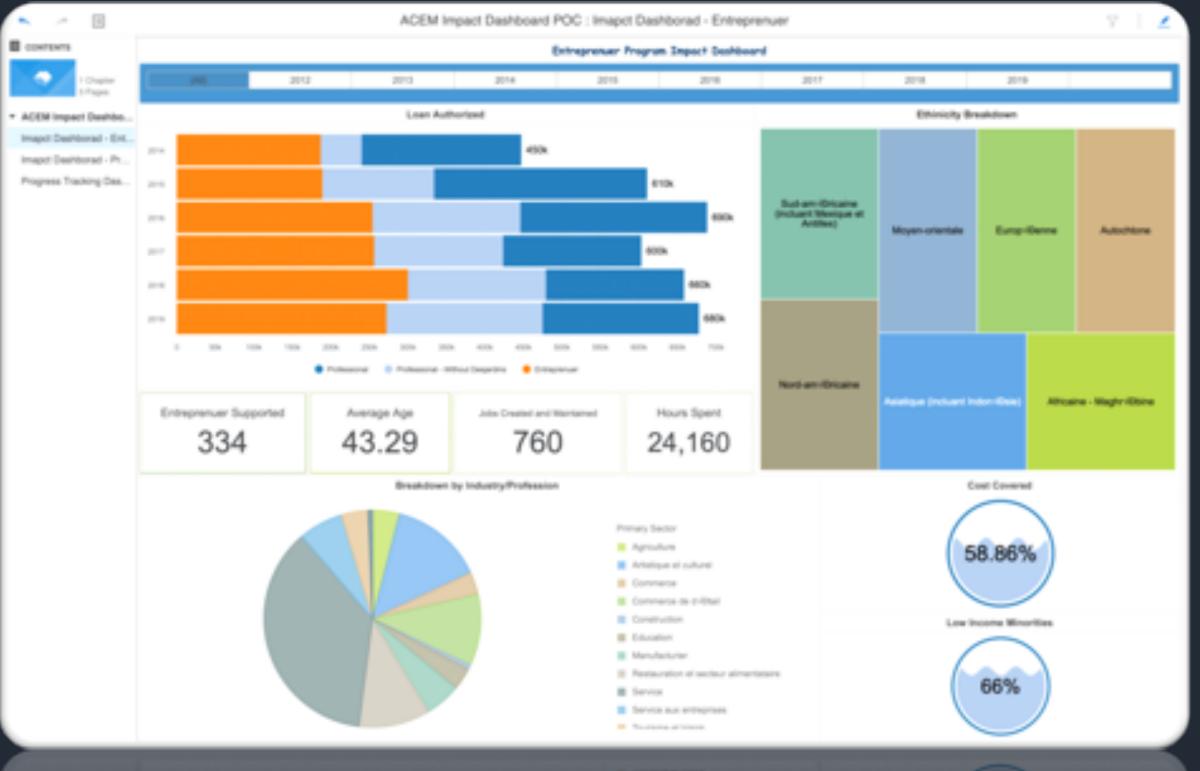


#### **Benefits**:

- Automated data cleaning and formatting during the python transform processes
- 2. Dashboard-ready data will be automatically saved to local drives
- 3. Dashboard Tool is linked to local drives and can instantly display refreshed data

# Demo & Dashboard

## Dashboard









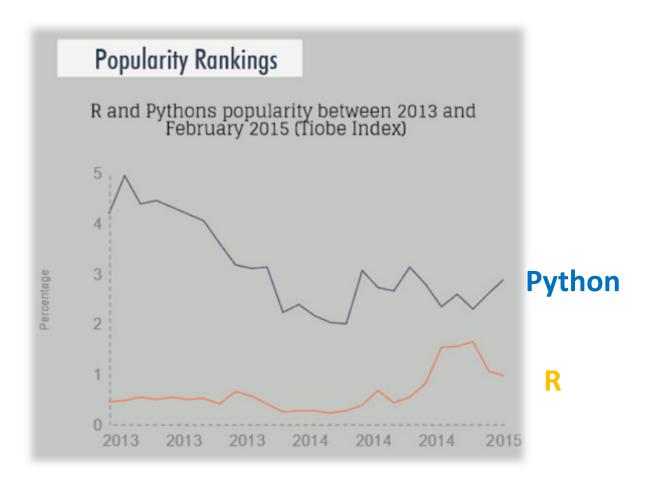
# Technical Evolution

## Adapting Advanced Analytical Tools: R & Python





Two of the most well-known and loved programming language in the analytics community \*



R focus on more user friendly data analysis, statistics and graphical models



**Purpose** 



Python emphasizes productivity and code readability

Primarily used by academic and research.

Statistical models can be written with only few lines



**Usability** 



Primarily used by programmers and developers Coding and debugging is easier (nicer "syntax")

Easy to use complex formulas in R, but have steep learning curve.

Have a collection of R functions that can be



Flexibility & Ease of Learning



Its readability and simplicity makes the learning more smooth and gradual.

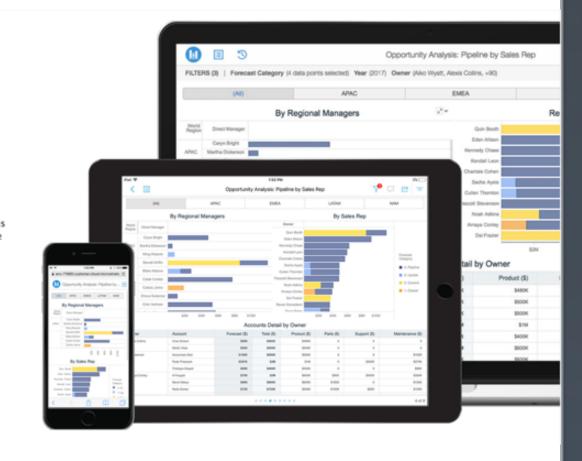
Have vast packages and libraries to be installed for deeper analysis

installed for deeper analysis

## **Enrolling to Premium**

## MicroStrategy<sup>®</sup>







#### Powerful Data Discovery

Explore your data using an intuitive toolset and built-in templates.



#### Compelling Data Visualizations

Leverage an extensive library using a drag-anddrop interface to quickly visualize data.



#### Intuitive Dossiers

Organize your dashboards, reports, and visualizations into chapters and pages.

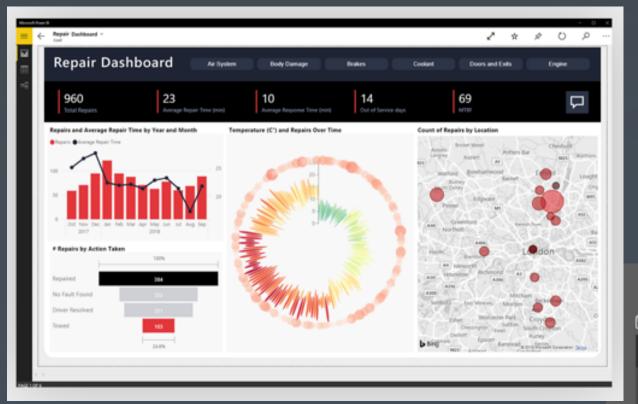
#### Go further with advanced analytics

The enterprise platform supports both web and mobile apps and comes with a suite of tools for easy administration, scalable in-memory technology, and enterprise security.

MicroStrategy doesn't publicly display its pricing information. Need to contact vendor for quote.

## Dashboard Tools





Power BI pricing

Analytics for every organization

Do you need self-service BI?

Do you need self-service BI?

Choose Power BI Pro

• Self-service and modern BI

• Collaboration, publishing, sharing, and ad hoc analysis

• Fully managed by Microsoft

Power BI Pro

\$12.80

Monthly price per user

Bo you need advanced analytics, big data support, on-premises, and cloud reporting?

Add Power BI Premium

• Enterprise BI, big data analytics, cloud and on-premises reporting

• Advanced administration and deployment controls

• Dedicated cloud compute and storage resources

• Allow any user to consume Power BI content.

Power BI Premium

\$6,393.60

Monthly price per dedicated cloud compute and storage resource

Get started >

Request a consultation >

Three Pricing Plans:

#### Power BI Desktop: FREE

data cleaning and preparation, custom visualizations and the ability to publish to the Power BI service.

Power BI Pro: \$12.8/user/month.

data collaboration, data governance, building dashboards with a 360-degree real-time view and the ability to publish reports anywhere.

(free trial for 60 days before purchasing the subscription.)

Power BI Premium: starts at \$6,000 a month per dedicated cloud compute and storage resource.

# Conclusion

## Our Solutions



#### **ACEM Project 1**

Representation and visualization of work done by ACEM into quantifiable societal impact. To attract and generate more donations.



#### Research & Analyze

Focusing on different user groups/investors perception

Benefits from seeing and evaluating overall higher level impact through relevant program



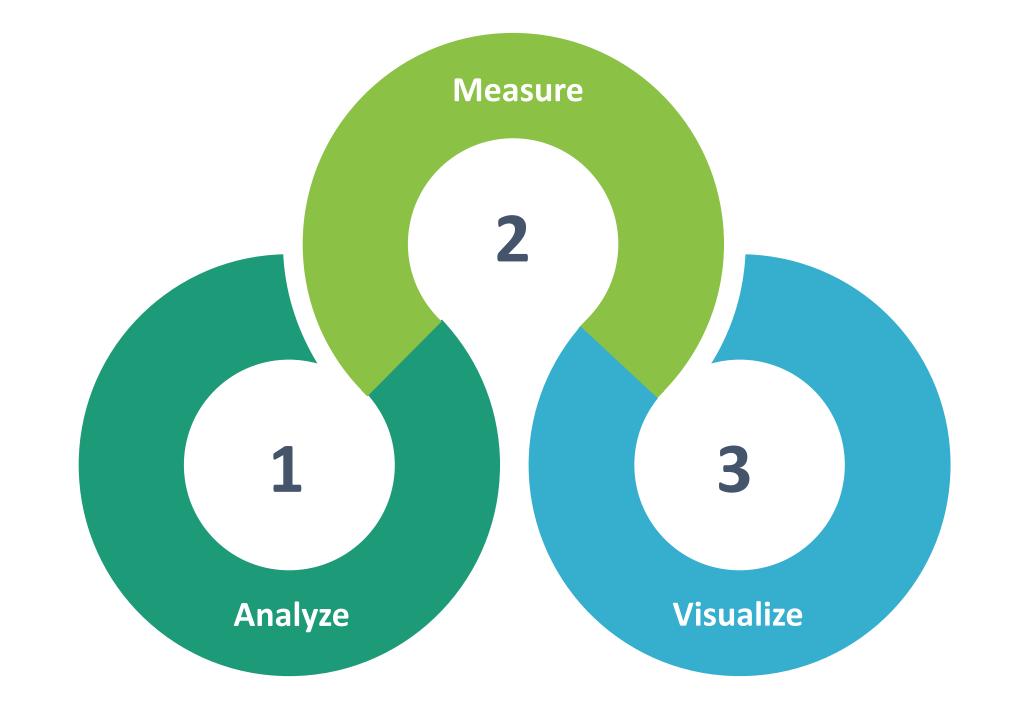
#### **Clean Data & Structured Metrics**

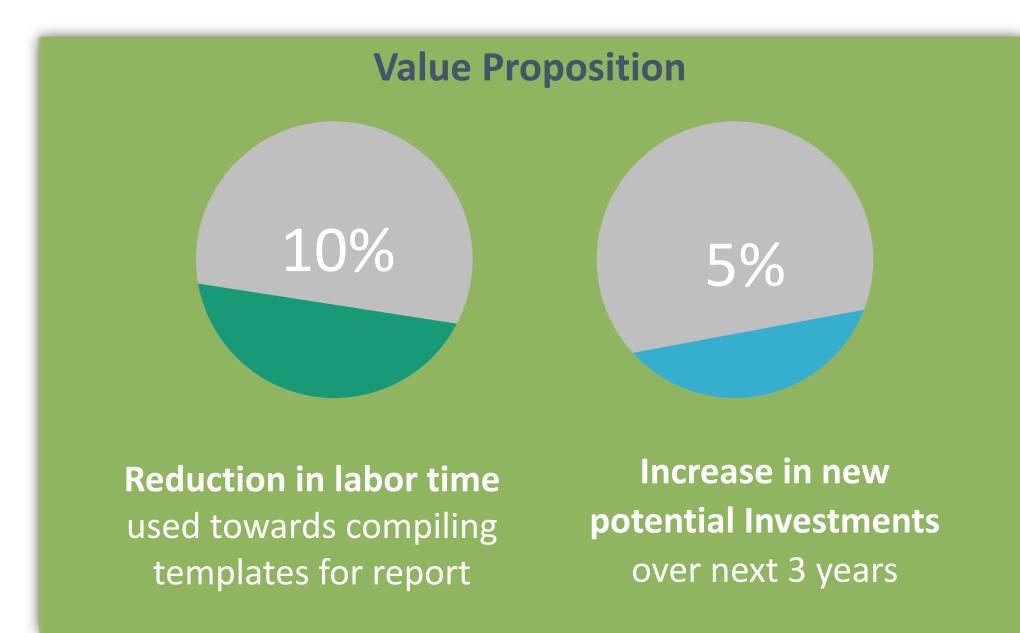
Tools that allows easier data cleaning and advanced analytics to be performed Better manage loans in a way to maximize social impact



#### **Dashboard**

Visualized metrics to communicate and showcase better convincible results to attract potential investors and donors



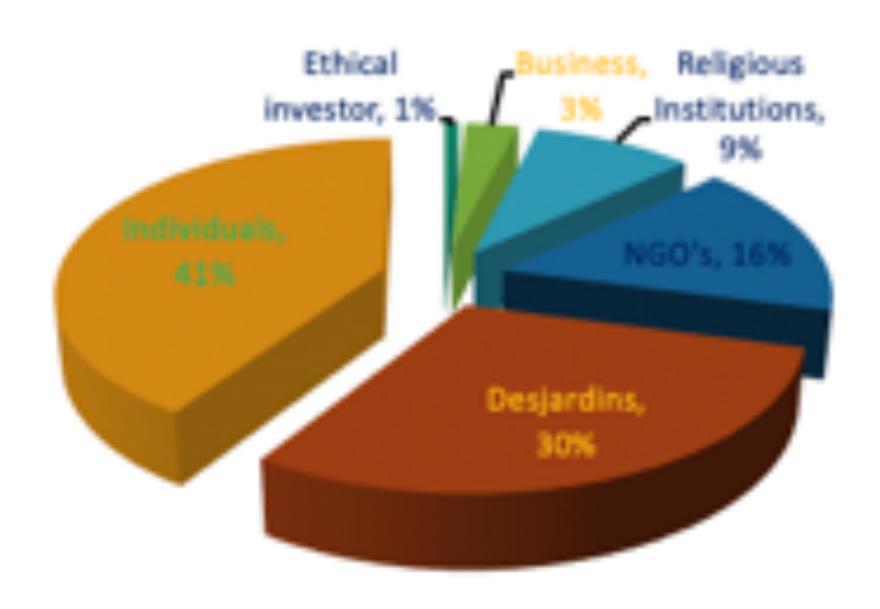


# Thank You!

**Q&A?** 

# Appendices

## Appendix 1: Investors/Donors Detailed Breakdown



#### **Categorization:**

- Individuals
- Commercial Business: Desjardins, Business
- Ethic Groups: Religious Institutions, Ethical Investors, NGOs

## Appendix 2: Python Tool - Instructions

\*Please refer "Python User Manual.docx" for more instructions and installation guide (Within Junkang's submission)



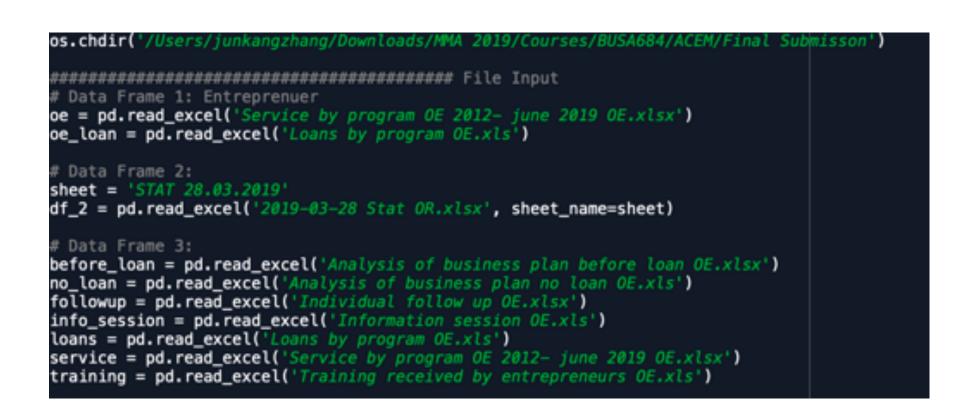
Step 2: Click
Python Run
Button

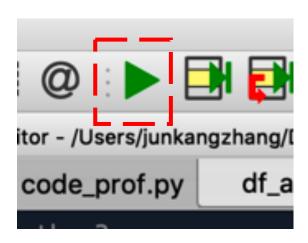
Step 3: Check New Data Files Generated

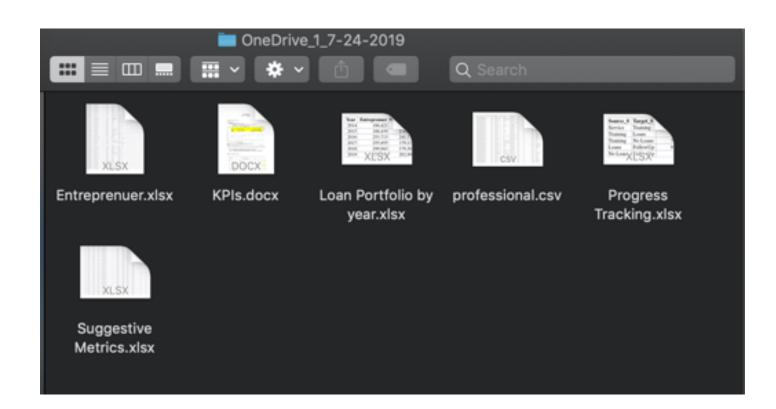










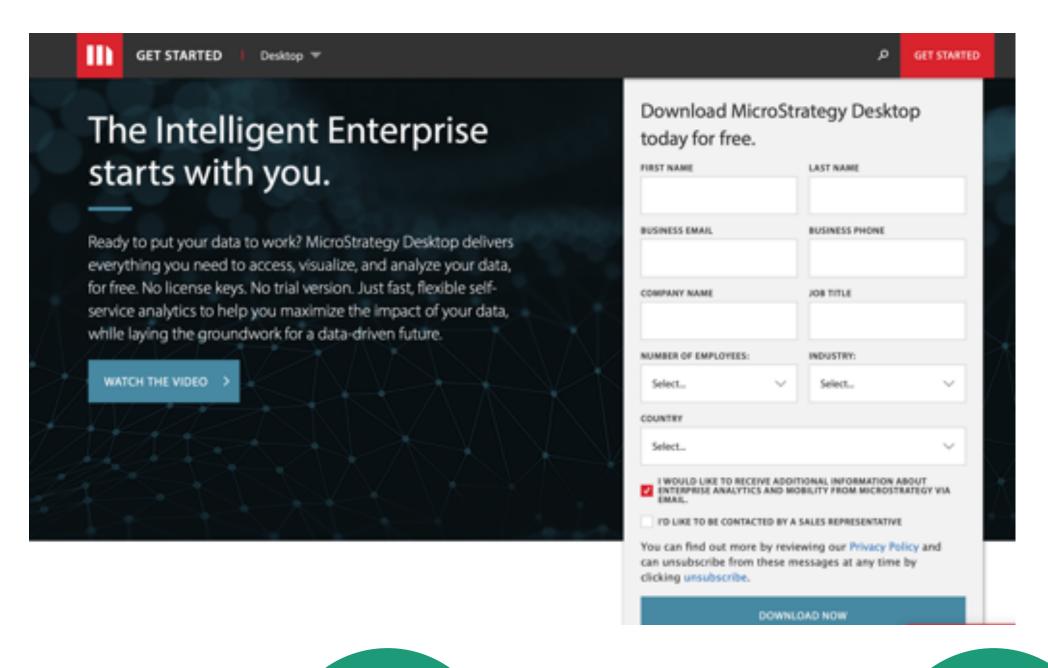


\* This section is the input section, it only takes a new user 5 minutes to set up the input directories

\* Simply click the "run" button to execute the program, no coding skills needed

\* New data files are automatically stored at the directory pre-specified for dashboard use

## Appendix 3: Preparing & Installing MicroStrategy Desktop





1

**Open Dashboard Template** 

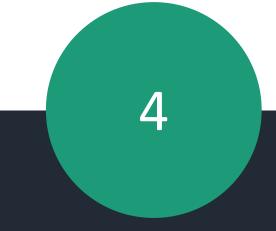
https://www.microstrategy.com/us/get-started/desktop

**Download Free Desktop** 

Open file:

ACEM Dashboard POC.mstr





→ DASHBOARD!

Duplicate as Metric

Create Attribute...

Create Links...

Number Format

Define Geography

Link to Other Dataset..

Amount of lo

Amount of pr

Amount of th

Avg (Age at

Avg (Numbe

Avg (Time be

#### **Import Relevant Files**

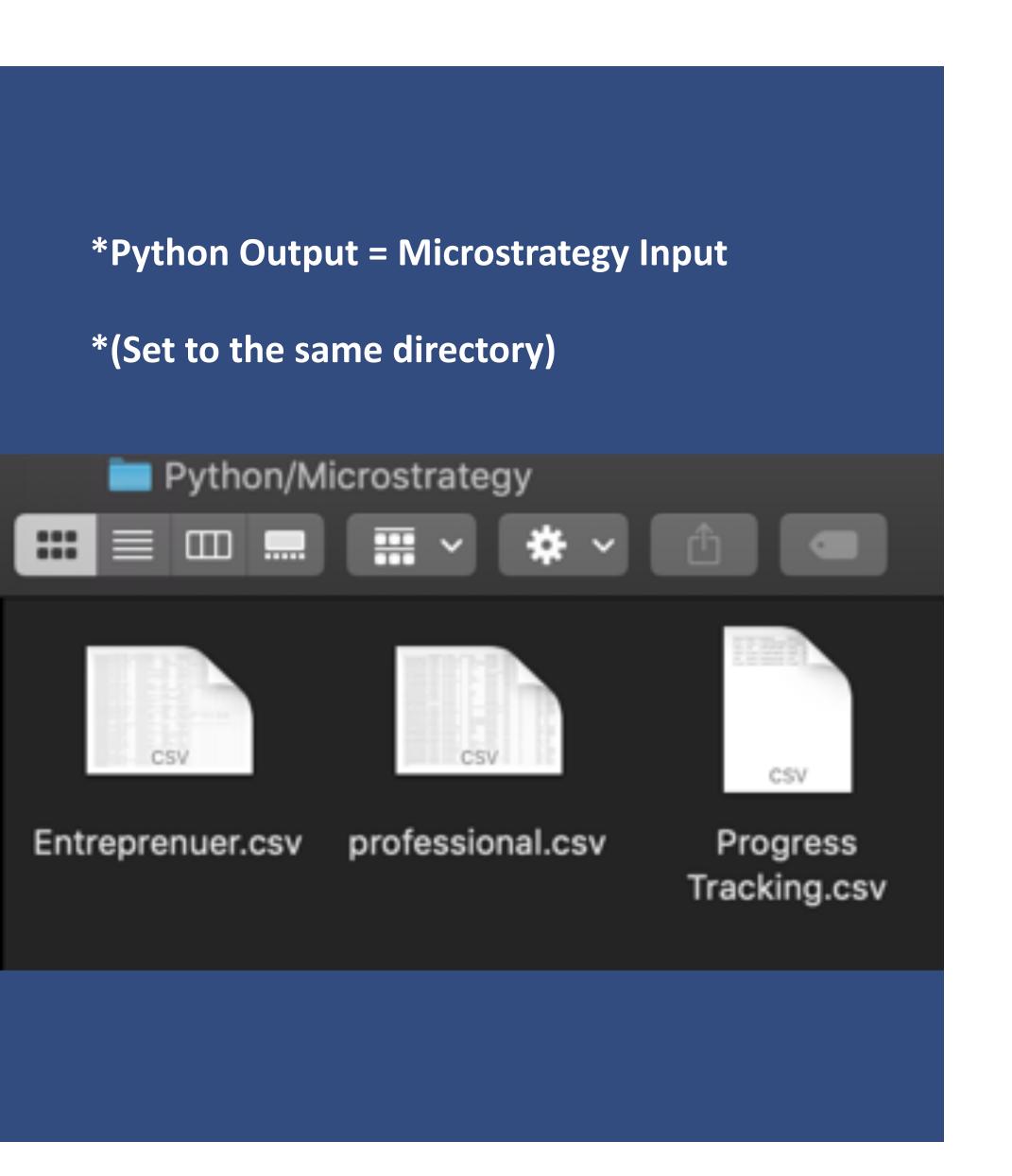
Import dataframe files and visual plugins.

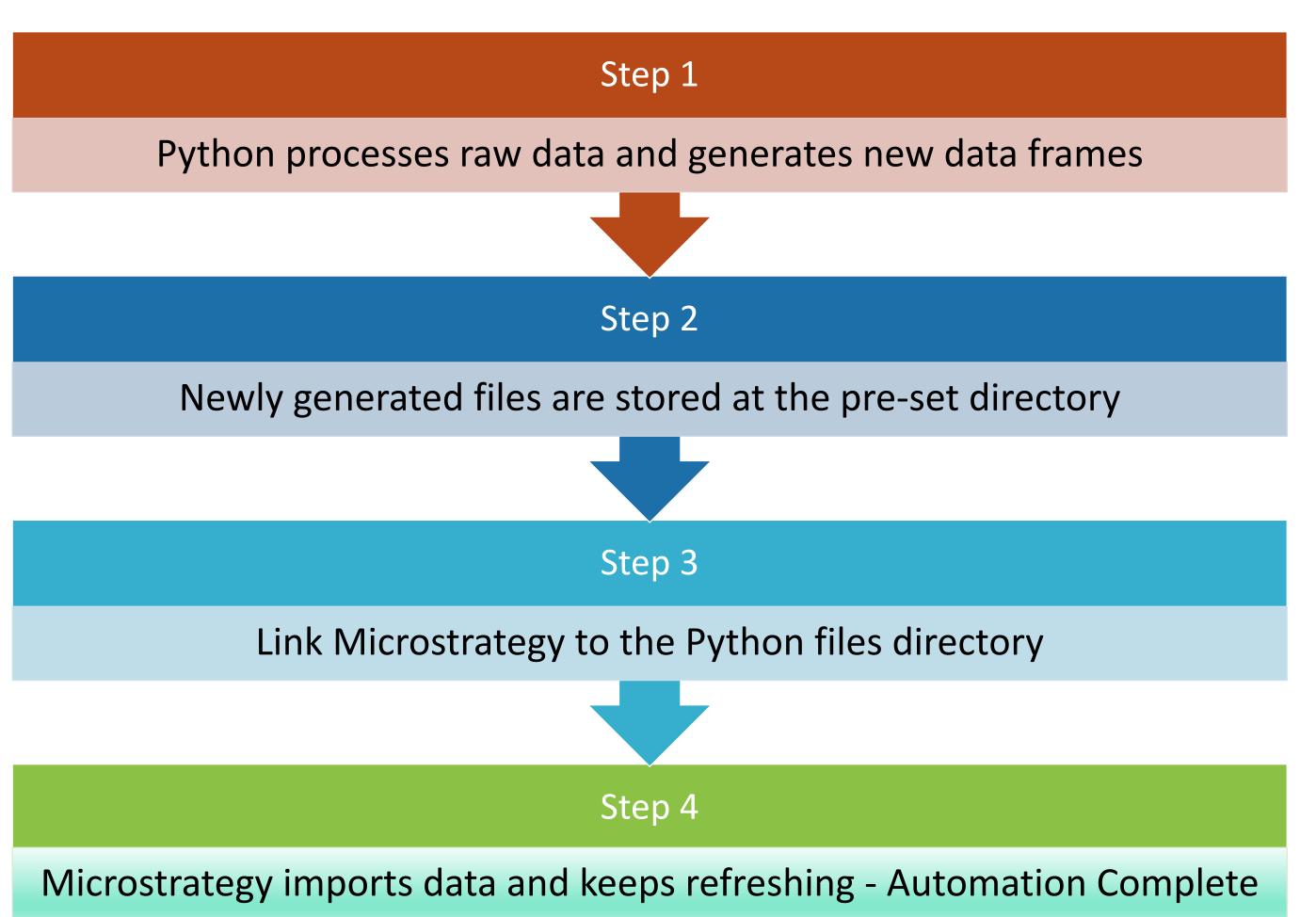
Format values where necessary

#### Link dataset

Link dataset by common field (i.e. Year) to enable dynamic filtering

## Appendix 4: Automation between Python and MicroStrategy





## Appendix 4: Predictive Model Instructions

Method	Mean Absolute Error
KNN	2.15179
Decision Tree	2.31866
Random Forest	2.86661

Three models were tested and KNN has the best performance

#### Future steps can be taken

- 1. Test out more machine learning techniques
- 2. Further automate the process3
- 3. Integrate it into the dashboard

#### Note:

These steps above can only be taken with the increase of data volume and data quality

#### **Note:**

Please see python file "predictive model.py" for more details on the algorithms