Statement of Work

Public Transportation Performance Analysis

Summary

IDA of TCCS commenced a project of parsing Canberra Light Rail GTFSR, comparing with static GTFS, Calculating Payment Mechanism Indexation and presenting results and analysis in local dashboard, in order to help CMET verify monthly payment statistics provided by Metro Operation companies. This statement of work provides details to outline project objectives, requirements, timelines, expectations, risks and acceptance.
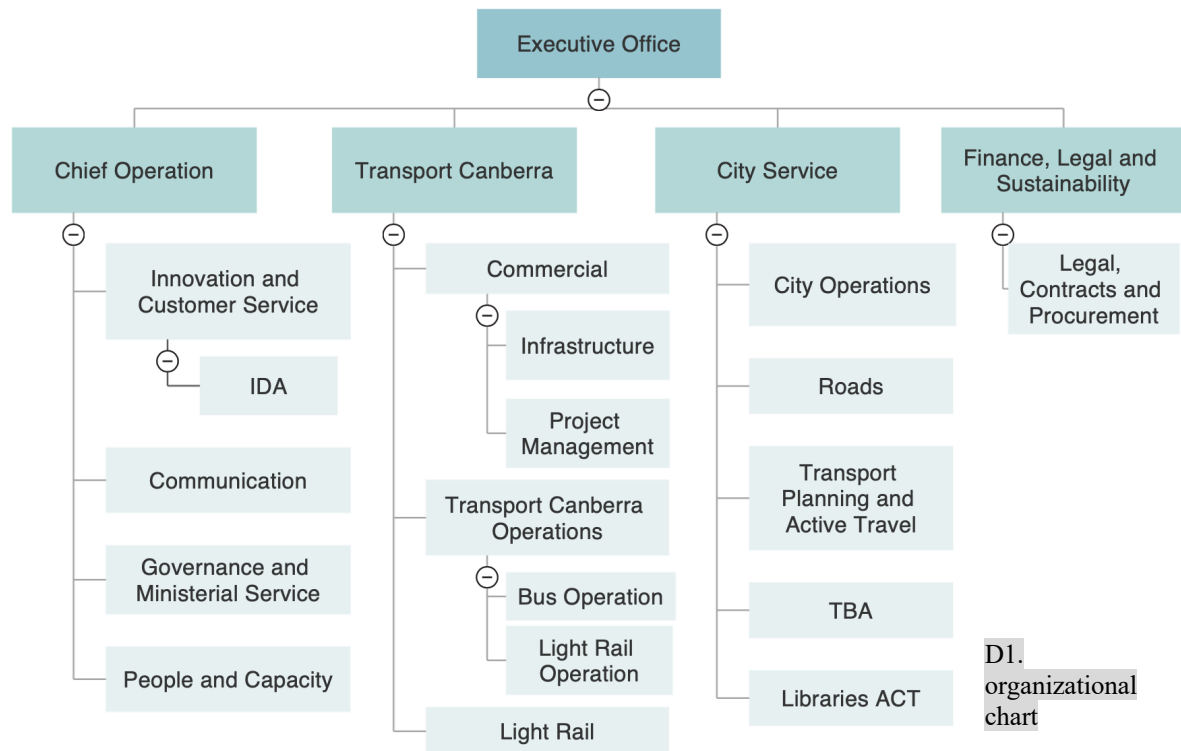
CONTENTS

DIAGRAMS LIST

1. Host Introduction

- **IDA and CMO**

The Innovation, Data and Analytics (IDA) Team is my main host, which is an enabling unit for reform and innovation across TCCS, which aims to continually improve services and implement innovative programs to deliver services and information to the Canberra Community.

The other associated stakeholder is CMO (Canberra Metro Operation), which is responsible for CMET, a group of companies that have been contracted by the ACT Government to operate and maintain the Light Rail network.

This organizational chart I draw below is from the information provided by TCCS, as I simplified few teams which are not strongly related to my project.

Executive Office

- **Chief Operation**
  - Innovation and Customer Service
    - IDA
  - Communication
  - Governance and Ministerial Service
  - People and Capacity
- **Transport Canberra**
  - Commercial
    - Infrastructure
    - Project Management
  - Transport Canberra Operations
    - Bus Operation
    - Light Rail Operation
  - Light Rail
- **City Service**
  - City Operations
  - Roads
  - Transport Planning and Active Travel
  - TBA
  - Libraries ACT
- **Finance, Legal and Sustainability**
  - Legal, Contracts and Procurement

D1. organizational chart

Deliverable or Task	IDA	Analyst	CMET	ICS	CMO	Totur	ANU
Phase 1							
Requirements elicitation	R	R	C	I	I	C	I
Payment Mechanism	C	R	R	C	C		I
GTFSR schema	A	R	C	I	C	I	I
Phase 2							
GTFSR Parsing	A	R	C	I	C	I	I
Static Schedule Comparison	A	R	A	C	I		I
Indexation Calculation	C	R	I	I	C	I	I
Phase 3							
Dashboard design	A	R	A	I	I		I
Dashboard implementation	A	R	A	I	I	I	I
Phase 4							
Test	A	R	A	I	C		I
Documentaiton	C	R	C	I	I	I	C

This RACI matrix illustrates the relationships between tasks of the projects and stakeholders.

D2. RACI

- **Interaction**

IDA is a new team across TCCS, devoting itself to supporting and enlightening analysis of data to provide innovative insight to decision making and reporting and help business units predict and plan for future trends. The interaction is across the whole TCCS, however especially closely associated with Customer Service and Transport Operations.

IDA has been continuously collecting customer passengers' feedbacks via phone, survey online and text messages, data of which is used to provide insight of service quality and

identify potential improvement for Transport Operation. IDA provides GTFSR data support to Transport Operation of light rail and buses. Moreover, data analysis can be utilized to outline principles in the future Planning and City Operations project. Below is the illustration of the interaction process of IDA.

D3. TCCS interaction
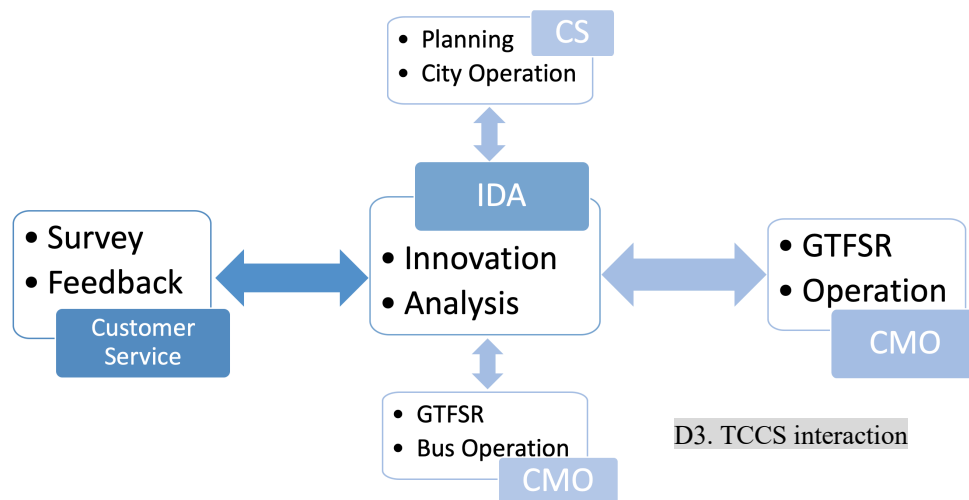
More specifically, the interaction process related to my project can be divided into two directions:

- One direction is from CMET GTFSR data (live operation of light rail geographic data) to IDA, which collect continuously updating GTFSR data and push to historical archive dataset of Open Data of Canberra. IDA supports the process of parsing, cleaning and updating the GTFSR, while CMET is responsible for providing accurate raw GTFSR data to reflect the current live performance of light rail operations.
- The other direction is started from IDA to CMO. In this process, IDA needs to provide analysis of light rail operations data to CMO, which can be used to track the performance of light rail and help verify the statistics provided by CMET. And in the long term, these analyses can be beneficial to improvement of light rail performance.

D4. IDA & CMO
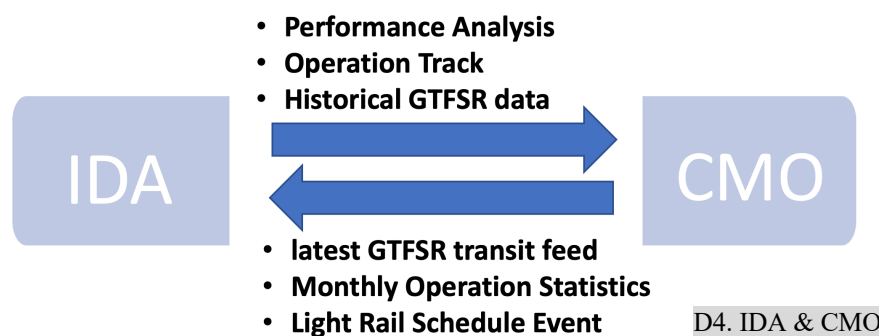
2. Vision and Objectives

- **Vision**

IDA commenced projects to enhance deliver an integrated, convenient, reliable and efficient transport system for Canberra by designing a more flexible and scalable platform that provides real time information including disruption notifications, integration with the new Light Rail and provide multi-modal journey options across Canberra.

Performance of light rail operations is the crucial part of public transportation, precisely track and analysis needs to be carried out. However, there is lack of proper process tackling this

task. Especial, with the help of automating data collection, benefit of transportation performance analysis can help decision making, issues prediction and future trend identification.

- **Project Objectives**

Goal: Timely Statistics Verification and Light Rail Performance Presentation

Current verification of Light Rail performance for Payment Contract lacks supplementary statistics, which hinders quicker responds and adjustment of service planning. Besides IDA, CSI and CMO are in the demand of statistics analysis to mechanism contract.

Objectives:

- CMO will use this dashboard as supplementary verification to monthly statistics provided by CMET, more frequently, timely and automatically.
- The performance analysis on the dashboard could facilitate associated teams across TCCS to identify and predict underlying issues of light rail operations, to improve service quality and deliver information to public.

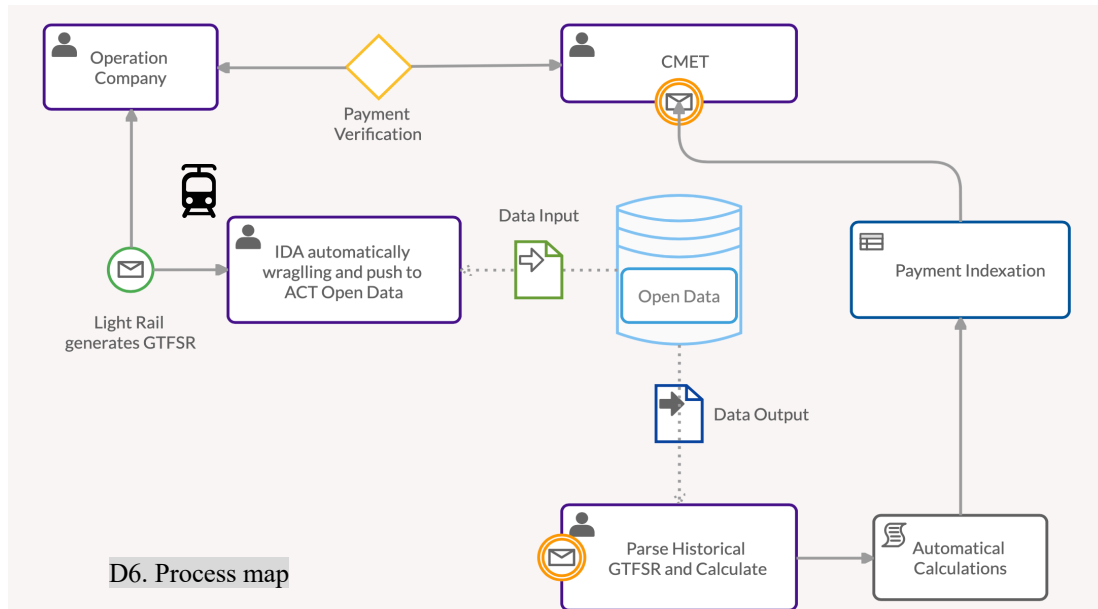Here is the Wireframe of the project schema, shown as below.
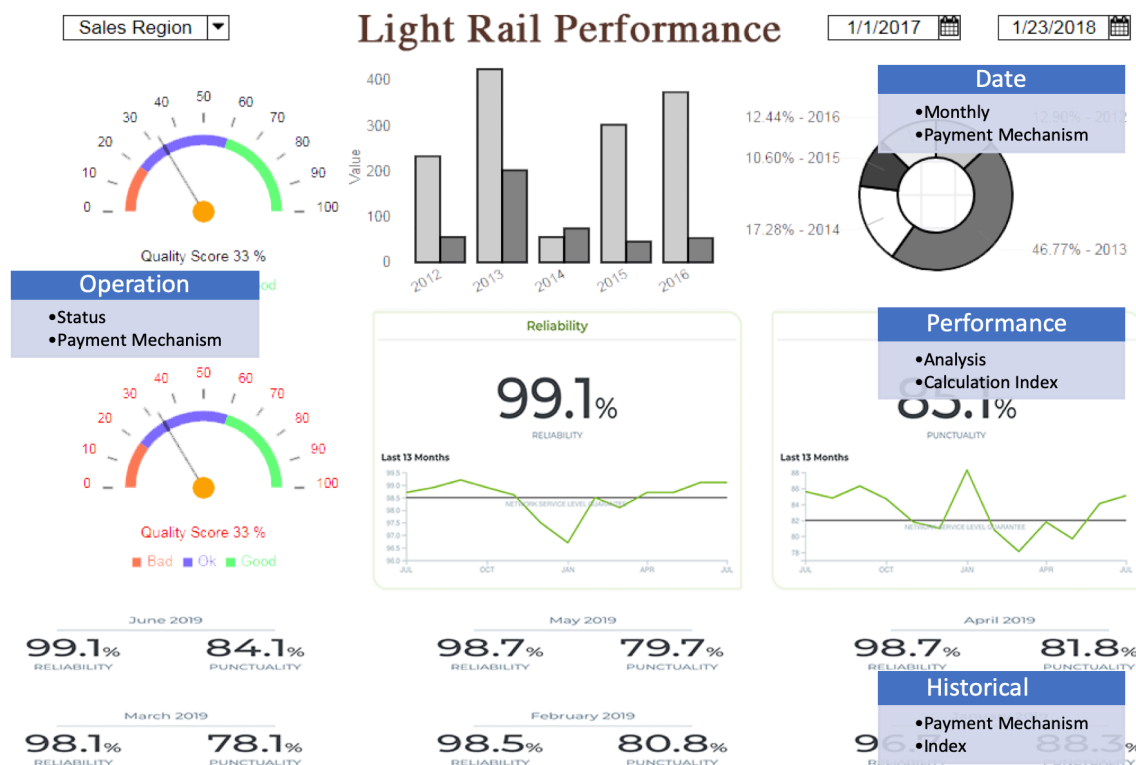
D5. Wireframe

3. Business Specification and User Requirements

- **Local dashboard to present Light rail performance**
  - **Functional Requirements:**
    1. Automatically GTFSR collection and parsing
    2. Identify Payment Contract events in parsed GTFSR
    3. PSA, ESA, LSA, MA, PCSR indexation and analysis
    4. Monthly performance statistics results
    5. Performance statistics results analysis

- **User Requirements:**
  1. Utilize dashboard results to verify CMET payment mechanism statistics
  2. Use light rail operation issues on dashboard to improve schedule of light rail
  3. Compare dashboard results with GTFSR data to identify manually errors during GTFSR generation process

This process map illustrates how activities flow across associated stakeholders.

D6. Process map

4. Milestones

Light Rail Performance Analysis	PROGRESS	START	END
Phase 1 Initiation and Requirements elicitation			
Start-up meeting and stakeholder contacts	100%	7/22/19	7/22/19
Roles , responsibilities requirements	100%	7/23/19	7/23/19
Agile KanScrum meeting schedule	100%	7/30/19	7/30/19
Project status and plans	100%	8/1/19	8/1/19
Phase 2 GTFSR schema and Payment Mechanism			
GTFRS complementation documents	100%	8/6/19	8/13/19
Light rail operation association logics	90%	8/6/19	9/3/19
Google transit feed standard library to parse GTFSR	100%	8/6/19	8/9/19
Payment Mechanism Contracts: PSA, ESA, LSA, MA, PCSR	90%	8/7/19	8/10/19
Payment indexation calculations logics	80%	8/8/19	9/3/19
Phase 3 GTFSR data parse and calculation			
Build test for implementation of parse and calculations	10%	8/20/19	9/3/19
Implement Python parser and calculator to calculate Light Rail	10%	8/20/19	9/19/19
Manually verify performance results with previous payment	0%	9/17/19	9/19/19
Keep code documents in accordance to Bitbucket repository	20%	9/9/19	9/19/19
Present preliminary analysis to stakeholders and modifications	0%	9/11/19	9/19/19
Phase 4 Dashboard design and implementation			
Design structure for dashboard	0%	9/24/19	9/26/19
Implement dashboard	0%	9/24/19	10/1/19
Check analysis accuracy and modidication	0%	10/1/19	10/3/19
Phase 4 Project closure			
Keep documentations in accordance with KanScrum and	0%	10/8/19	10/17/19
Complete project requirements checklist, to achieve delivery	0%	10/10/19	10/19/19

D7. Timeline 1

Phase 1   **Project Initiation and Requirements elicitation**

Purpose   To explore the problem faced by business stakeholder, understand their business context, make agreements on deliverables, project deadlines

Activities

- Project start-up meeting and stakeholder contacts
- Roles and responsibilities and identify requirements and establish delivery requirements
- Make weekly agile KanScrum meeting schedule with supervisor to track progress and issues
- Present project status and plans for following week

Deliverables

- Acknowledge and agree project requirements
- Agree project commencement

Phase 2   **GTFSR schema and Payment Mechanism**

Purpose   To explore data format to solve the business problem, understand their business system, databases, external data sources.

Activities

- Read GTFRS complementation documents to understand GTFSR data structure and identify light rail operation plots
- Utilize google transit feed standard library to parse GTFSR data to verify data schema understandingAccess to data sets
- Read Payment Mechanism Contracts documents to understand payment conditions measurement: Passenger Service Availability (PSA), Early Services Adjustment (ESA), Late Services Adjustment (LSA), Monthly Availability (MA), The Partially Completed Services Result (PCSR).
- Outline calculations logics between Payment Measurement index and GTFSR data

Deliverables

- Parsing data storage structure design

Phase 3 **GTFSR data parse and calculation**

Purpose Data parsing and wrangling will be performed to solve the business problem

Activities

- Build test to check following implementation of parse and calculations
- Implement Python parser and calculator to calculate Light Rail performance
- Manually verify performance results with previous payment statistics
- Keep code documents in accordance to Bitbucket repository analysis
- Present preliminary analysis to stakeholders

Deliverables

- Data analysis technical report
- Data analysis code/program

Phase 4 **Dashboard design and implementation**

Purpose outline dashboard design and structure

Activities

- Design structure for dashboard

Deliverables
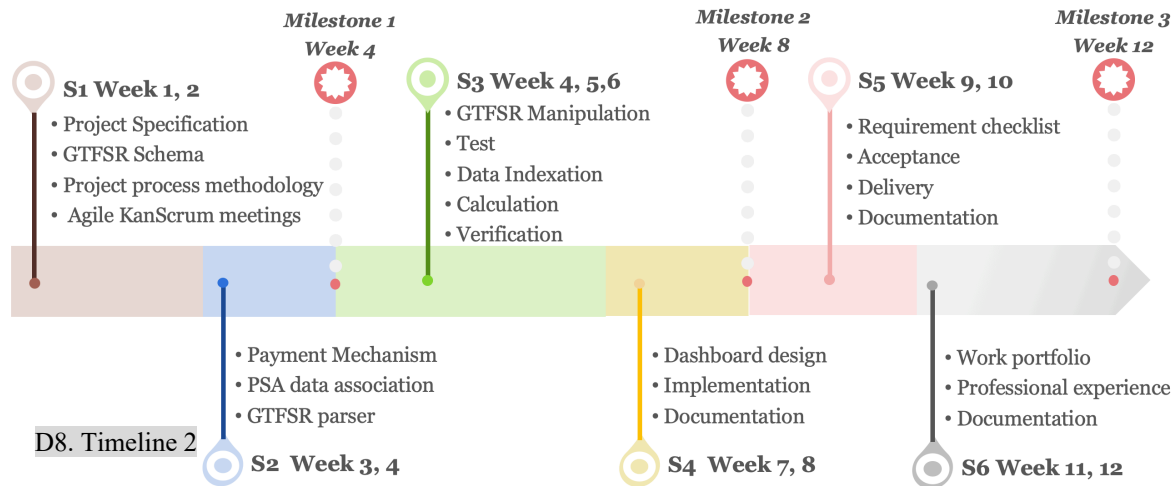
- Dashboard wireframe

Phase 5 **Project closure**

Purpose outline dashboard design and structure

Activities

- Keep documentations in accordance with KanScrum and check Objective and Confluence Page

Deliverables

- Complete project requirements checklist, to achieve delivery requirements
- Work portfolio package

Milestone 1 — Week 4
Milestone 2 — Week 8
Milestone 3 — Week 12

S1 Week 1, 2
- Project Specification
- GTFSR Schema
- Project process methodology
- Agile KanScrum meetings

S3 Week 4, 5,6
- GTFSR Manipulation
- Test
- Data Indexation
- Calculation
- Verification

S5 Week 9, 10
- Requirement checklist
- Acceptance
- Delivery
- Documentation

D8. Timeline 2

S2 Week 3, 4
- Payment Mechanism
- PSA data association
- GTFSR parser

S4 Week 7, 8
- Dashboard design
- Implementation
- Documentation

S6 Week 11, 12
- Work portfolio
- Professional experience
- Documentation

5. Acceptance Criteria and deliverables

- Local version dashboard of light rail performance
- Dash board presents daily performance and monthly payment indexation: PSA, ESA, LSA, MA, PCS
- Based on the quality of Historical Archive Data of DTFSR, identify monthly PCS, MA
- Documentations as guidance for dashboard users and illustrate associations between GTFSR and payment indexation

6. Resources and Risks

- **GTFSR parsing**:
  1. GTFRS complementation documents illustrates the data format, header, field and domain, cardinality and categories, which can be used to understand schema of light rail performance data
  2. Transit feed in open source library is implemented in python2, which would define the compatibility for interface with dashboard. This could bring risks if these codes need to be compatible with python3 or have more confidential requirement for open source library
  3. IDA has implemented python script to collect updating GTFSR and transform and push it into ACT Open Data Set. And I identified possible errors (primary key set) could influence the accuracy of arrival time records after mid night trams. This features could influence the payment indexation calculations based on this data.

- **Payment Indexation Calculation:**
  1. CMET report monthly payment indexation and statistics to CMO, which can be used as investigation to calculation from historical GTFSR
  2. Payment Mechanism documentations contain definitions for PSA, ESA, LSA, MA, PCSR, which identifies different service application for GTFSR parsing and calculation
  3. Previous payment statistics can be used during implementation process to help address errors

- **Dashboard design and implementation**:
  1. Use previous Power Bi project to keep in accordance to dashboard design

2. Use Plotly template to implement dashboard, which needs to be adjusted to satisfy customer requirements

- **Documentations:**

  Use Confluence and Bitbucket to maintain documentations during project process, in accordance with KanScrum process requirements

Risk Description	Risk Level	Treatment Option	Mitigation Strategy	Note
Inaccuracy of GTFSR	Medium	Check Data features	Use primary key and identification to mitigate	
Special Event	Low	Read possible record from GTFSR	Manual import of operation data	
Static GTFS	Medium	Update parsing frequently	Sent frequent request to check last modified timestamp	
GTFSR generation errors	Medium	Adjust according to statistics from CMET	Adjust PCS results according to CMET report	
Parsed data overflow	Low	Write to Jason files and re-run on PC	Write parsed data into Json files and operate on PC	

D9. Risks

7. Tooling

- **Parsing Tooling**

Here is the Paired Comparison for Parsing Tooling decision making, since the parsing for GTFSR data could use open source library, which is more time efficient. This decision would be consistent with python script used to push latest GTFSR into ACT Open Data. Besides requirements for confidentiality and security for parsing process is relatively low.

Parser Comparison					
	Feature 1	Feature 2	Feature 3	Ranking	Notes
Google Transit Feed	Python2	Attribute format	Open Source	1	Low requirement confidentiality
Modification Google Library	Python2	Standard Json	Secure	2	Specific Parsing
Customized parser	Python3	Standard Json	Secure	3	Specific Parsing

D10. Parser Comparison

- **Calculation Tooling**

Since the payment mechanism indexation is fixed according Contract Agreement conditions, the standard python statistic library could not address the specific features: PSA, ESA, LSA, MA, PCSR.

Calculation Comparison					
	Feature 1	Feature 2	Feature 3	Ranking	Notes
Python Statistics	Python3	Standard statistical	Modify parsed	2	Standard library

	Feature 1	Feature 2	Feature 3		Notes
Functional Calculator	Python2,3	Specific statistical	Json format	1	Customized calculation

<p align="right">D11. Calculation Comparison</p>

- **Dashboard Tooling**

Considering compatibility and integrity with parser and calculator, adapting python plotly would address the delivery quality and improve user experience.

Dashboard Comparison					
	Feature 1	**Feature 2**	**Feature 3**	**Rank ing**	**Notes**
Python Plotly	Python2, 3	Flexible design	Integrated with parser	2	Compatible with parser
Power Bi	IDA	Standard visualization	Separate from data	1	Separate from parser

<p align="right">D12. Dashboard Comparison</p>