

Proof of Concept & Discovery Phase for Data Analytics Platform

The client is a large, healthcare-focused strategic media planning and buying group of several agencies. The group offers media planning for different channels by analyzing existing data available within each agency. They wanted to develop an analytics platform to improve executive decision-making across the agency.

To help our client be strategic about creating such a platform, we engaged in a **Proof of Concept (POC) and Discovery Phase** to do the following: (1) identify how feasible it would be to use existing agency data to effectively analyze and measure performance and trends, and (2) understand and document the business and technical requirements for developing an analytics platform.

Business Benefits:

- We recommended a big data-based solution with a clear path for execution. This approach will help our client do the following:
 - Attain business vision of scalability, maintainability and sustenance
 - Manage data traits like volume, variety and historical data persistence
 - Accommodate the future vision of predictive modeling, taking into account the growth of data volume

Engagement Deliverables

2. Users
The main user for this platform would be Data Analysts from different agencies.

3. Primary Goals

- 3.1. Loading historical data
- 3.2. Create mechanisms / process for loading ongoing data
- 3.3. Benchmarking
- 3.4. Reporting
- 3.5. Predictive Analytics

4. Functional Requirements

REQ #	Requirement	Module	Agency
FR_001	The system shall authenticate the user trying to log in using Active Directory	Authentication	All
FR_002	The user should be able to assign users to the following roles: 1. Analyst (agency level) 2. Analyst (Group level) 3. Agency Admin 4. Group Admin	Authorization	All
FR_003	The system allows access to features in the platform based on their roles	Authorization	All
FR_004	The feature access rules shall be configurable	Authorization	All
FR_005	User should be able to create data access rules for based on user roles, the agency that they belong to and access	Authorization	All

5. Operational Requirements

REQ #	Requirement	Module	Agency
	into the system	Onboarding	All
	cy's name. The used after the	Onboarding	All
	ncy, this will that agency, rit into the	Onboarding	All
	's name. The ated after the	Onboarding	All
	's name, this ampaigns	Onboarding	All
	id for a client	Onboarding	All
	The status of	Onboarding	All
	's name. The ated after the	Onboarding	All
	ids, but this will aigns	Onboarding	All

The collage also includes several data visualization dashboards showing various charts and maps, such as a line chart with a trend line, a bar chart, a map of the United States, and a horizontal bar chart.

Key Highlights

Proof of Concept (2 months with 2 resources)

During this phase, we performed following activities:

- Defined success criteria for POC
- Extracted data from multiple data sources
- Cleaned up data and ingested it into the data warehouse
- Classified data as qualitative and quantitative
- Validated data integration business rules to secure and isolate data access
- Performed data analysis across agencies (brands) and created custom reports

Discovery Phase (6 weeks with 1 onsite and 2 offshore resources)

- Gathered and documented key business and technical requirements
- Understood different stakeholders' visions and business goals
- Identified data sources for the analytics platform's data warehouse
- Examined and determined expected data volumes and the type of data to be analyzed
- Defined a recommended solution based on the aforementioned activities
- Defined the recommended technology stack

Proposed Technologies

- Kafka Connect, Kafka (with Zookeeper), Cassandra, Spark (with Mesos), Ubuntu, Tableau, Scala, Python, Java