

Tipo : Enunciado de laboratorio
Capítulo : Exploración del Big Data
Duración : 40 minutos

I. OBJETIVO

Conocer cómo aplicar una metodología de proyecto de Big Data.

II. REQUISITOS

Los siguientes elementos de software son necesarios para la realización del laboratorio:

- Microsoft Office 2013 Professional o mayor
- Conexión a internet

III. EJECUCIÓN DEL LABORATORIO

1. Su empresa requiere realizar un proyecto de Big Data, con este fin es necesario presentar un Project Chapter. Para esto deberá crear un documento en Word con la siguiente estructura:

Project Charter

Conocimientos del negocio

- ¿Quién es el cliente? ¿En qué dominio comercial se encuentra el cliente?
- ¿Qué problemas comerciales estamos tratando de resolver?

Alcance

- ¿Qué soluciones de ciencia de datos estamos tratando de construir?
- ¿Qué es lo que se va hacer?
- ¿Cómo va a ser consumido por el cliente?

Personal

- Quienes están en el proyecto:
 - Equipo:
 - Líder de Proyecto
 - Especialista Big Data
 - Analista de Datos
 - Analista de Negocio
 - Cliente:
 - Empresa

Métricas

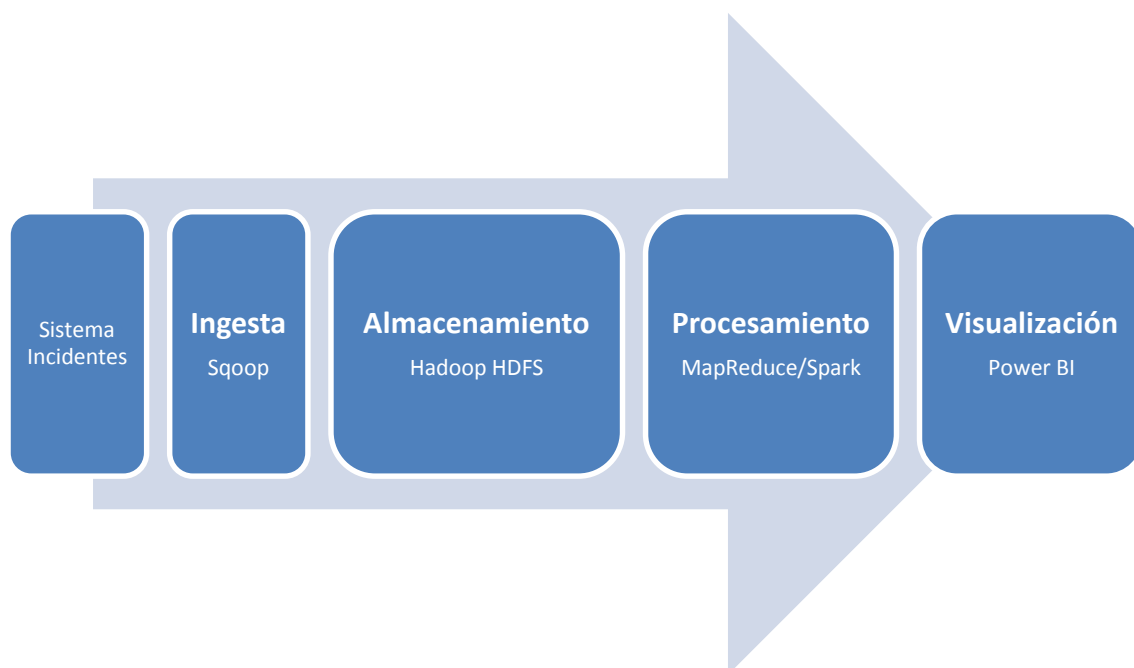
- ¿Cuáles son los objetivos cualitativos? (por ejemplo, reducir la rotación de usuarios)
- ¿Qué es una métrica cuantificable (por ejemplo, reducir la fracción de usuarios con inactividad de 4 semanas)?
- Cuantifique qué mejoras en los valores de las métricas son útiles para el escenario del cliente (por ejemplo, reduzca la fracción de usuarios con inactividad de 4 semanas en un 20%)
- ¿Cuál es el valor de referencia (actual) de la métrica? (por ejemplo, la fracción actual de usuarios con inactividad de 4 semanas = 60%)

Plan

- Fases (hitos), línea de tiempo, breve descripción de lo que haremos en cada fase.
- Por ejemplo: Se han definido cuatro fases para el proyecto:
 - Ingesta de Datos:
 - Almacenamiento de Datos:
 - Procesamiento de Datos:
 - Análisis de Datos:

Arquitectura

- Datos
 - ¿Qué datos esperamos? Datos brutos en las fuentes de datos del cliente (por ejemplo, archivos de texto, SQL, Hadoop, etc.)
- Herramientas y recursos de almacenamiento / análisis de datos que se usarán en la solución
- Diagrama de la solución. Por ejemplo:



Comunicación

- ¿Cómo nos mantendremos en contacto? ¿Reuniones semanales?
- ¿Quiénes son las personas de contacto en ambos lados?

Diccionario de Datos

Por ejemplo:

ITEM	CAMPO	TIPO	VALORES	DESCRIPCIÓN	DATASET
1	CodigoIncidente	Int	1001	Código del incidente reportado por el ciudadano	Incidentes
2	Incidente	String	Foco intermitente	Descripción del incidente reportado por el ciudadano	Incidentes
3	CodigoSensor	String	S001	Código del incidente generado por el sensor	Sensor
4	Sensor	String	Sensor de ruido	Descripción del incidente reportado por el sensor	Sensor

Fuentes de Datos

Por ejemplo:

ITEM	DATA SET NAME	ORIGINAL LOCATION	DESTINATION LOCATION
1	Sensor	IOT	HDFS
2	Incidente	Sistema de Incidentes	HDFS