



**Tecnológico
de Monterrey**

Análisis de datos Huawei

Primer Avance Proyecto

Laboratorio de diseño y optimización de operaciones

Profesora: Ana Luisa Masetto Herrera

Gerardo Gutiérrez Aguirre	A01368146
Juan Ramón Rodríguez Márquez	A01365002
José Manuel González Guadarrama	A01364681
Kevin Alberto Sorzano Mongel	A01368603
Mario Naime Monroy	A01365605
José Alberto Naime Monroy	A01365892

Etapas 1

1) Descripción de la situación actual:

Huawei es una empresa china de telefonía móvil que actualmente se encuentra como el tercer más grande vendedor de smartphones en México, ocupando aproximadamente un 14% de la participación en el mercado. Su situación en el mundo se complicó por los bloqueos tecnológicos de Estados Unidos en el gobierno pasado, provocando una caída en sus ventas que hasta este punto no se ha podido recuperar. El panorama para la compañía está lleno de retos entre los que destacan reducir la diferencia con sus competidores, un tema que con la llegada de sus nuevos aparatos y sus estrategias en ventas desean lograr lo más pronto posible.

La importancia de este tipo de proyectos recae en obtener una ventaja competitiva a través del análisis de datos internos para su correcta interpretación y una posterior planeación estratégica a través de los mismos.

2) Descripción de la Problemática

El negocio de la venta de celulares ha sido una montaña rusa para Huawei, pues empezó como una fuerte amenaza para sus competidores situándose en 2019 como el segundo mayor vendedor de celulares en México con un 16% de participación en el mercado (sólo superado por Samsung), sin embargo las restricciones del gobierno de EUA derivaron en una baja de ventas para el año 2020 y una caída del segundo al tercer lugar en México. El panorama actual no luce alentador es por eso que se busca a través de un análisis de datos, explotar los puntos estratégicos aprovechando los recursos internos para poder figurar de nuevo en los dos primeros puestos de ventas de celular.

Como ingenieros industriales podemos constatar que este tipo de problemáticas son importantes para nuestra carrera ya que muchos de los diferentes empleos que están relacionados con ella tienen que ver con data science y la forma en la que resolvemos problemas y digerimos la información obtenida. Muchos de estos empleos requieren que podamos ubicar, filtrar y encontrar datos específicos y que su rápida localización sea indispensable para poder facilitar el proceso de búsqueda, por lo que es indispensable aplicarlo de la mejor manera.

3) Entender y describir la problemática (en términos de ciencia de datos, tipo de tarea, tipo de datos, etc.).

Enfrentamos un problema de regresión debido a que manejamos datos cuantitativos y queremos obtener el mejor dato numérico para la correcta predicción de la

demanda y buscamos responder a la pregunta “¿Cuál es la forma óptima de trabajo para la organización ya que los costos sean mínimos con la mayor ganancia?”. Mencionado lo anterior, a continuación mostramos material que nos ayudará a entender mejor la problemática de la empresa y con ello poder definir de mejor manera los objetivos.

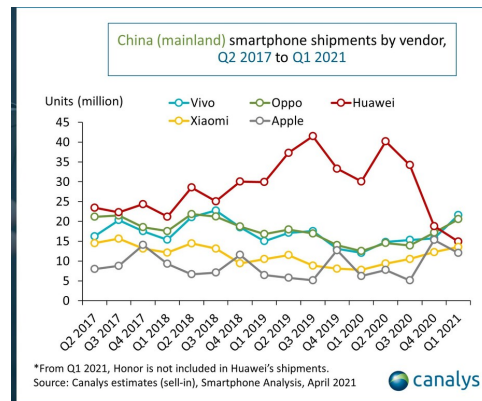


Fig 1: Gráfica de envíos de celulares por proveedores (cuartos de año).

Las estadísticas muestran que Huawei ha pasado de tener 41% del mercado total de smartphones en China a un 16% en los últimos tres periodos mostrados en la tabla 1. Vivo y Oppo, han superado claramente a Huawei para este año 2021 y la tendencia indica que Xiaomi también lo hará.

Huawei shipments continue to hemorrhage

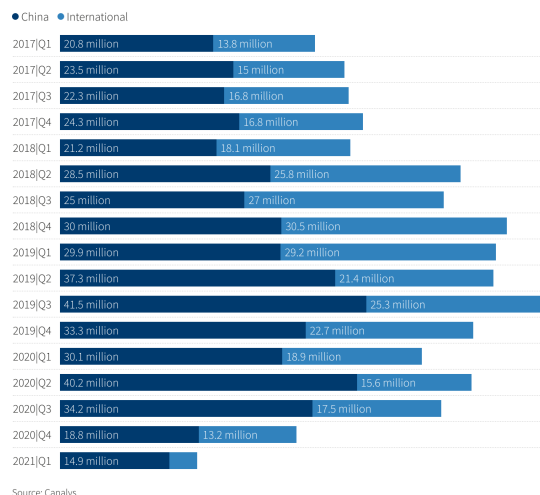


Fig 2: Gráfica de envíos de celulares Huawei

En la tabla 2, podemos observar que el decrecimiento es inminente no solo en China. Sino que a nivel internacional Huawei se está derrumbando. La empresa ha sido claramente afectada por diferentes factores, donde sobresale el veto impuesto por la administración de Trump y mantenido por la nueva de Biden. La empresa está enfrentando una etapa nunca antes vivida, por ende en términos económicos podríamos asumir que se encuentran en una crisis.

4) Plasmar los objetivos.

Como objetivos personales definimos los siguientes. Los dividimos en primarios y secundarios con el fin de secuenciar el proceso de trabajo para asegurar el correcto funcionamiento del mismo. Nuestro objetivo primario se basa en la correcta realización del proyecto haciendo un buen sistema para la predicción de la demanda por medio de la metodología CRISP DM que sea capaz de resolver o disminuir el impacto de las adversidades que se presentan en la empresa. Los objetivos secundarios los planteamos basándonos en las expectativas de los medios del proceso, pensando de forma sistémica y secuencial tenemos los siguientes puntos para cada una de las etapas:

- **Etapas 1 – Comprensión del Negocio:** descubrir desde el principio factores importantes que pueden influir el resultado del proyecto.
- **Etapas 2 – Comprensión de los datos:** familiarización con los datos. Describir y clasificar datos crudos, detectar problemas de calidad.
- **Etapas 3 – Preparación de los datos:** filtración de ellos. Realizar una limpieza óptima de los datos. Aplicación de ingeniería estadística.
- **Etapas 4 – Modelado:** codificación y análisis de datos. Planteamiento de modelos.
- **Etapas 5 – Evaluación:** evaluación de los modelos propuestos anteriormente y análisis de resultados. Conclusiones.
- **Etapas 6 – Implantación:** aplicación del proyecto y presentación final.

El fin general es que el proyecto sea realizado de manera correcta utilizando las herramientas de programación y procesamiento de datos vistas en clase.

5) Estructurar el proyecto y hacer un plan preliminar.

Con base en la organización de las actividades del proyecto, se hizo la siguiente distribución y asignación de tareas con tiempos establecidos para posteriormente realizar en diagrama de Gantt y poder visualizarlo de mejor manera:

Al ser un equipo de trabajo de seis integrantes, designamos un líder para cada etapa según nuestros rasgos y habilidades.

Las organización de equipo se encuentra en el apartado de “Anexos”

Debido a la extensión del diagrama no pudimos obtener una buena imagen del mismo, sin embargo, en el link a continuación se puede encontrar:

https://docs.google.com/spreadsheets/d/1EoAdx8eBGwgOO9F_BIm3qpjoA02rlqyl/e/dit?usp=sharing&oid=114674023006189029027&rtpof=true&sd=true

Etapas 2

1) Descripción de los datos crudos.

Los datos con los que vamos a trabajar vienen en un archivo de Excel este archivo de Excel cuenta con 369,618 renglones ,también tiene 14 columnas el archivo cuenta con 14 variables que son:punto de venta,fecha,mes,año,núm ventas, sku, marca, gamma, costo promedio, zona, estado, ciudad, latitud y longitud. Se cuentan con 1902 puntos de venta en 9 zonas distintas, 237 ciudades y existen tres tipos de gama: alta, baja y media. Para mayor duda

2. El Diccionario de variables se encuentran en el apartado de “Anexos”

2) Problemas de calidad con los datos

A continuación se listan los problemas que se encontraron a la hora de observar los datos:

- Hay un problema con los meses, debido a que algunos están escritos con números y otros están escritos con abreviación.
- En la columna de año, año está escrito como "anio", además hay otro problema debido a que toda la columna debería estar escrita con números de 4 cifras como "2018" y hay algunos años que están marcados como "18".
- Con la variable de zona hay un problema con la ortografía ya que algunos datos están escritos con mayúsculas y otros con minúsculas, además hay unas zonas que están repetidas pero escritas de diferentes maneras
- En la columna de marca también existe un problema de ortografía con las mayúsculas y minúsculas
- Dentro la columna de estados hay datos erróneos ya que dentro de esta hay registros de ciudades como "Metepéc" y "La Paz", cuando estas no deberían estar ahí porque son ciudades

3. Las preguntas de cierre de documento se encuentran en el apartado de “Anexos”

Referencias

- Carlos Fernández De Lara. (2021). *Vamos a regresar al trono, pero nos tomará tiempo: Huawei*. 2021, de Forbes Sitio web:
<https://www.forbes.com.mx/tecnologia-vamos-regresar-trono-nos-tomara-tiempo-huawei/>
- Pastor, J. (2021, 30 abril). *Debate de Huawei en China: el veto provoca la caída en picado de sus ventas incluso en su país de origen*. Xataka.
<https://www.xataka.com/moviles/debate-huawei-china-veto-provoca-caida-picado-su-s-ventas-incluso-su-pais-origen>
- BBC News Mundo. (2021, 5 marzo). *Huawei: cómo la falta de acceso a componentes está asfixiando al gigante tecnológico chino*.
<https://www.bbc.com/mundo/noticias-56216518>

Anexos

1. Organización del equipo

Etapas 1 - Comprensión del negocio. Líder: Gerardo Gutierrez

- Descripción de la situación actual. (Duración aproximada: 1 día)
- Describir la problemática (en términos del negocio). (Duración aproximada: 1 día)
- Describir la problemática (en términos de ciencia de datos). (Duración aproximada: 1 día)
- Plasmar los objetivos. (Duración aproximada: 1 día)
- Estructurar el proyecto y hacer un plan preliminar. (Duración aproximada: 1 día)

Etapas 2 - Comprensión de los datos. Líder: Jose Alberto Naime

- Describir los datos crudos. (Duración aproximada: 1 día)
- Detectar problemas de calidad. (Duración aproximada: 1 día)

Etapas 3 - Preparación de los datos. Líder: Jose Manuel Gonzalez

- Limpieza de datos. (Duración aproximada: 5 días)
- Análisis exploratorio. (Duración aproximada: 7 días)

- Seleccionar y construir variables para la etapa de modelado. (Duración aproximada: 7 días)

Etapas 4 - Modelado. Líder: Mario Naime

- Promedios móviles. (Duración aproximada: 9 días)
- Aprendizaje de máquina. (Duración aproximada: 9 días)

Etapas 5 - Evaluación (Resultados). Líder: Kevin Sorzano

- Construcción de gráficas. (Duración aproximada: 6 días)
- Conclusiones. (Duración aproximada: 5 días)

Etapas 6 - Implantación. Líder: Juan Rodríguez

- Reporte final. (Duración aproximada: 8 días)
- Presentación ejecutiva. (Duración aproximada: 7 días)

2. Diccionario de variables

- **Punto_de_venta:** Es el espacio en el que una empresa establece contacto con su posible cliente potencial.
- **Fecha:** Es un indicador a través del tiempo en que ocurre o se hace algo.
- **Mes:** Periodo que transcurre entre un día y el de igual fecha del mes siguiente.
- **Año:** Periodo a contar de doce meses, a contar desde un día cualquiera.
- **Num_ventas:** Es el conjunto de ventas de productos o servicios que ha realizado una empresa en un periodo determinado.
- **Sku:** Los códigos sku son elementos fundamentales encargados de llevar el control y gestionar el stock en el almacén.
- **Marca:** Es un nombre, término, señal, un símbolo o una combinación de algunos de ellos que identifica.
- **Gama:** Es un índice que clasifica a los teléfonos con respecto a sus atributos.
- **Estado:** Estado de la república mexicana.
- **Costo_promedio:** Es el costo por cada unidad dentro de una producción
- **Zona:** Es la extensión de terreno cuyos límites están determinados por razones económicas, políticas, entre otras.
- **Ciudad:** Es una población donde habita un conjunto de personas que se dedican a actividades industriales y comerciales.
- **Longitud:** Distancia angular de un punto de la superficie terrestre, según los meridianos.
- **Latitud:** Distancia angular que hay desde un punto de la superficie de la Tierra hasta el paralelo del Ecuador.

3. Preguntas de cierre de documento

¿Cómo van con respecto a la limpieza de datos? Vamos avanzando fluidamente y el trabajo en equipo nos está haciendo filtrar los datos con unas estrategias que planteamos con una lluvia de ideas que tuvimos en nuestra primera sesión como equipo el lunes 30 de agosto. Concluimos que vamos bien.

¿Será necesario agendar alguna asesoría por equipo? Creemos que si es necesario ya que debido a nuestra falta de experiencia en esta rama de la carrera podríamos omitir sin darnos cuentas filtros y pasos en el proceso para un proyecto apropiadamente.

¿Qué etapas y actividades se deben de realizar en los siguientes entregables?

- Etapa 2 - Comprensión de los datos, describir los datos crudos, detectar problemas de calidad.
- Etapa 3 - Preparación de los datos, limpieza de datos, análisis exploratorio, seleccionar y construir variables para la etapa de modelado.
- Etapa 4 - Modelado, promedios móviles, aprendizaje de máquina.
- Etapa 5 - Evaluación (Resultados), construcción de gráficas, conclusiones.
- Etapa 6 - Implantación: Reporte final, presentación ejecutiva.