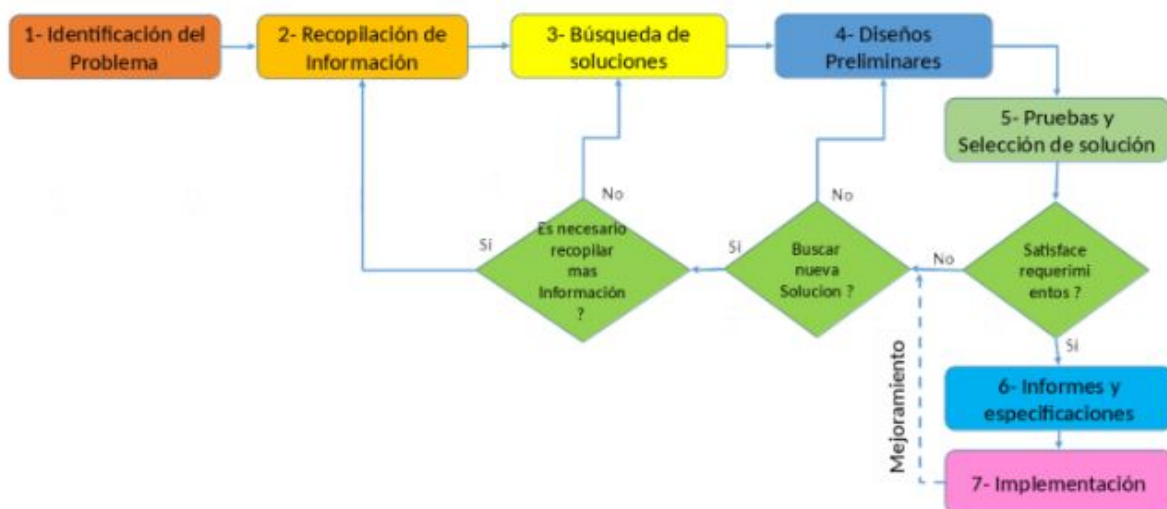


Christian Flor - A00355624, Carlos Restrepo - A00355028, Daniel Fernández - A00354694

Desarrollo de soluciones.

Se utiliza el Método de Ingeniería para desarrollar una solución a la problemática planteada posteriormente, siguiendo un enfoque sistemático de acuerdo con la situación problemática.

Basado en la descripción del Método de Ingeniería del libro "Introducción a la Ingeniería" de Paul Wright, presenta el siguiente diagrama de flujo, cuyos pasos seguiremos en el desarrollo de la solución.



FASE 1: IDENTIFICACIÓN DEL PROBLEMA.

Descripción del contexto problemático (causas y síntomas):

La oficina de estadísticas de transporte (BTS por sus siglas en inglés) hace parte del departamento de transporte de los estados unidos. Esta se encarga de compilar, analizar y hacer accesible la información sobre los distintos tipos de transporte que operan en dicho país. Entre los múltiples servicios que ofrece la agencia se encuentra la biblioteca nacional de transporte y la base de datos TranStats.

Según la BTS “ La Biblioteca Nacional de Transporte (NTL) de la Oficina es el permanente hogar de acceso público para publicaciones de investigación de toda la comunidad del transporte”. Esto es, la biblioteca nacional de transporte es un repositorio en línea en donde se guardan bases de datos e informes sobre transporte. Así mismo, TranStats es una base de datos que incluye estadísticas de aviación, carretera, peatones, entre otros.

Dentro de las bases de datos que incluye TranStats se encuentra “on time”. Una base de datos de desempeño en el tiempo de las aerolíneas, que incluye el desempeño a tiempo de cada vuelo, aerolínea y aeropuerto en los Estados Unidos.

Por otro lado, teniendo en cuenta los retrasos que sufren los usuarios de las aerolíneas para llegar a sus compromisos u obligaciones debido a los vuelos atrasados, se realizará una aplicación que haciendo uso de la probabilidad y de la información presentada por la base de datos “on time” le ofrezca a los usuarios la posibilidad de saber que tan probable es que se atrase en un vuelo que vaya a tomar.

Identificación de los síntomas y necesidades de la situación problemática

- El retraso de los vuelos puede interferir con los planes y obligaciones del viajero.
- Se requiere hacer uso de un mapa para que el usuario vea la información de manera más interactiva.
- Se requiere que los usuarios tengan información sobre cuál es la probabilidad de que su vuelo se retrase.

Definición del problema

Comúnmente las aerolíneas dan información a sus usuarios del tiempo que dura un vuelo de una ciudad origen a una ciudad destino, pero en esta información las aerolíneas no tienen en cuenta los posibles retrasos que pueden sufrir los vuelos.

FASE 2: RECOPIACIÓN DE LA INFORMACIÓN NECESARIA:

Para tener una claridad total en los conceptos involucrados, se realiza una búsqueda de las definiciones de los términos relacionados con el problema planteado. Es importante realizar esta búsqueda en fuentes reconocidas y confiables para saber qué elementos son parte del problema y cuáles no.

Además de las búsquedas web se le preguntó a un profesor del curso de Inferencia Estadística de la Universidad ICESI acerca de qué recomendaciones nos podía dar para escoger las variables que les podíamos sacar más provecho y así estimar las probabilidades correspondientes para cada variable.

Referencias.

- [1] "2017-Jan-OnTimeFlightData-USA," data.World, 2017. [Online]. Available: <https://data.world/hoytick/2017-jan-ontimeflightdata-usa>. [Accessed: 29-jan-2020]
- [2] "AirOnTime87to12: Airline On-Time Performance Data," Microsoft, 2019. [Online]. Available: <https://docs.microsoft.com/es-es/machine-learning-server/r-reference/revoscaler/airontime87to12>. [Accessed: 29-jan-2020]
- [3] "GMAP.NET BEGINNERS TUTORIAL MAPS MARKERS, POLYGONS AND ROUTES," independent-software, 2016. [Online]. Available: <http://www.independent-software.com/gmap-net-beginners-tutorial-maps-markers-polygons-routes-updated-for-vs2015-and-gmap1-7.html>. [Accessed: 27-jan-2020]
- [4] "Distribución de probabilidad", Wikipedia, 2016. [En línea]. Disponible en: https://es.wikipedia.org/wiki/Distribuci%C3%B3n_de_probabilidad. [Accedido: 30-ene-2020]
- [5] "El teorema del límite central: las medias de muestras grandes y aleatorias son aproximadamente normales", Minitab 18, 2019.. [En línea]. Disponible en: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/data-concepts/about-the-central-limit-theorem/>. [Accedido: 30-ene-2020]

Marco teórico

De acuerdo a la base de datos sugerida por nuestro mentores llamada "2017-Jan-OnTimeFlightData-USA" el cual tiene 450.017 datos y 66 variables se investigó que describe cada variable en un sitio web de microsoft docs el cual nos amplia la información de los campos, por lo consiguiente sabremos qué variables nos importan en esta gran base de datos para poder abordar nuestra problemática ya planteada.

Estado del arte.

Debido a que la solución debemos de implementarla con el recurso de GMaps.Net, realizamos el tutorial propuesto por nuestros profesores del curso el cual nos ayudó a comprender y apropiarnos más acerca de este recurso.

Nuestros compañeros del curso proyecto integrador I de la Universidad ICESI también están trabajando con el mismo entorno de desarrollo y herramienta (GMap.Net) pero cada grupo tiene que abordar diferentes problemáticas.

FASE 3: BÚSQUEDA DE SOLUCIONES CREATIVAS

Para este paso, como ya sabemos que contamos con suficientes datos para analizar la base de datos, se procede a la búsqueda de métodos para calcular la probabilidad de retraso:

★ Ideas para calcular la probabilidad de retraso de los vuelos

Alternativa 1. Distribución de probabilidad aleatoria.

Tomando el concepto de la distribución de probabilidad de una variable aleatoria el cual es una función que asigna a cada suceso definido sobre la variable la probabilidad de que dicho suceso ocurra y enlazandolo con nuestra problemática, se asemeja que nuestra variable (diferencia de tiempo) sería nuestra variable aleatoria y corresponde a una variable discreta de valores enteros por lo que frecuentemente toma valores enteros.

Alternativa 2. Teorema del límite central.

Utilizar el teorema del límite central el cual nos indica que si el tamaño de la muestra es lo suficientemente grande, la distribución de las medias sigue aproximadamente una distribución normal. Por lo tanto, al volverse una distribución normal la variable se volvería continua y tendría que tomar valores dentro un intervalo dado.

Alternativa 3. Organizar dataset.

Organizar la base de datos en sus respectivas columnas para posteriormente, filtrar solamente por ciudades y por el tiempo de retraso. Por último se agrupan los destinos con datos iguales sacando un promedio y posteriormente calcular la probabilidad de demora.

Year	Quarter	Month	DayofMonth	DayOfWeek	FlightDate
2017	1	1	17	2	17/01/2017
2017	1	1	18	3	18/01/2017
2017	1	1	19	4	19/01/2017
2017	1	1	20	5	20/01/2017
2017	1	1	21	6	21/01/2017
2017	1	1	22	7	22/01/2017
2017	1	1	23	1	23/01/2017

★ Ideas para la creación de la aplicación

Debido a que el entorno de desarrollo en el cual nos estamos profundizando es .NET utilizaremos GMap.NET ya que este nos permite crear mapas, marcadores y polígonos en su aplicación .NET / C#.

En el curso brindado de GMap.NET nos proporcionan una lista de algunos proveedores de mapas:

Alternativa A. BingMapProvider.

BingMap es una web de mapas creada por Microsoft para su buscador Bing.

Alternativa B. GoogleMapProvider.

Google Map es un servidor de aplicaciones de mapas en la web que pertenece a Alphabet Inc.

Alternativa C. YahooMapProvider.

Yahoo! Maps es un portal gratuito de mapas en línea proporcionado por Yahoo!

FASE 4: TRANSICIÓN DE LA FORMULACIÓN DE IDEAS A LOS DISEÑOS PRELIMINARES

En este paso descartamos las opciones inviables.

★ Ideas para calcular la probabilidad de retraso de los vuelos

Alternativa 1. Distribución de probabilidad aleatoria.

- Es una buena alternativa a tener en cuenta pero debido a que la variable es discreta, se deben de analizar con 3 tipos de distribuciones las cuales son:
 - Distribución binomial (eventos independientes).
 - Distribución de Poisson (eventos independientes).
 - Distribución hipergeométrica (eventos dependientes).

Por lo que tendríamos que analizar cada distribución y esto tomaría mucho tiempo.

Por lo cual queda descartada esta alternativa 1.

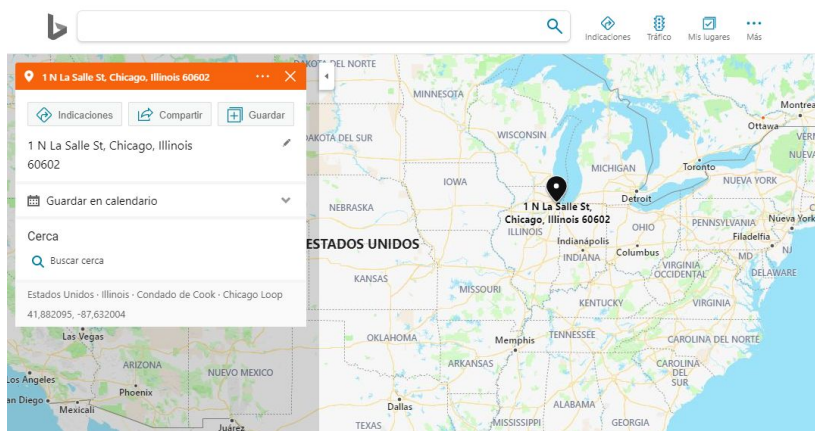
Alternativa 2. Teorema del límite central.

-Dicha alternativa habla de que si tomamos una muestra lo suficientemente grande y si se desconoce cómo se distribuye la variable, se puede tomar como una distribución normal. Además, la media poblacional y la media muestral serán iguales

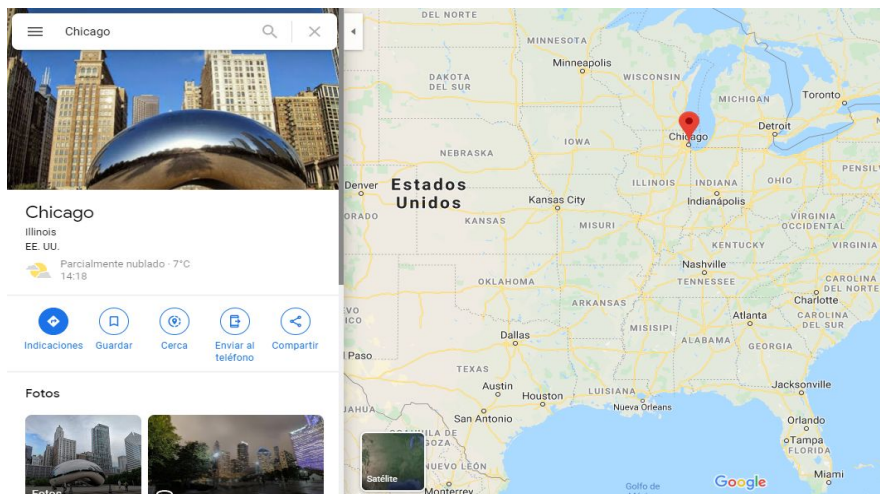
Al convertirse en una distribución normal los valores de la variable aleatoria estarán dentro de un rango determinado. Por lo cual esta alternativa la podemos tomar.

★ Ideas para la creación de la aplicación

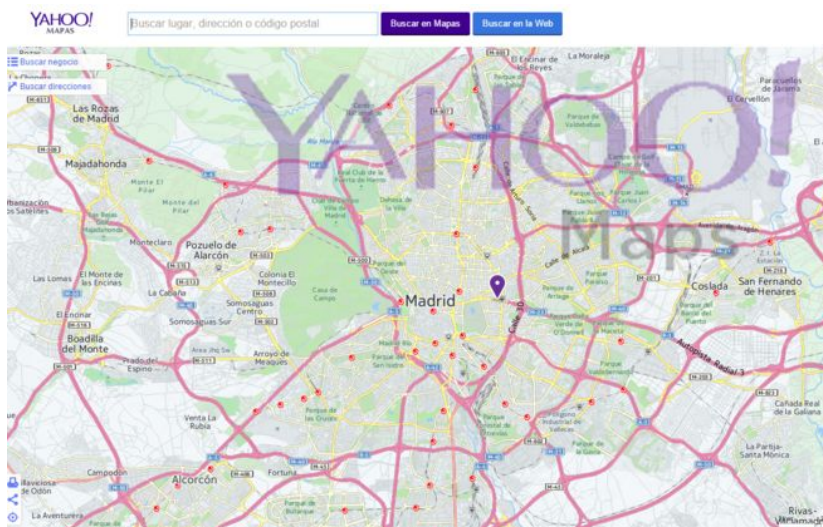
Alternativa A. BingMapProvider



Alternativa B. GoogleMapProvider



Alternativa C. YahooMapProvider



FASE 5: EVALUACIÓN Y SELECCIÓN DE LA MEJOR SOLUCIÓN

Se deben definir los criterios que permitirán evaluar las alternativas de solución para poder elegir la solución que mejor satisfaga las necesidades del problema. Los criterios que se eligieron en este caso son los que se enumeran a continuación. Al lado de cada uno se ha establecido un valor numérico para establecer un peso.

Criterios para la evaluación de las ideas propuestas para calcular la probabilidad de retraso de los vuelo.

Nivel	1	2	3
Criterio 1A: Facilidad de implementación	Baja: no se requieren más conocimientos.	Media: se necesita al menos un nuevo conocimiento.	Alta: se necesitan nuevos conocimientos.
Criterio 1B:	Dos días o menos	Entre tres y cuatro	Más de cuatro

Tiempo de implementación		días. (inclusivo)	días.
Criterio 1C: Uso de herramientas adicionales	No hay que usar más herramientas	Se usan máximo dos herramientas adicionales	Se usan más de dos herramientas adicionales.

*Si se tiene una alternativa con una puntuación baja será escogida como la mejor solución.

Evaluando con los criterios anteriores las alternativas que se mantienen, obtenemos la siguiente tabla:

Evaluación de ideas para calcular la probabilidad de retraso de los vuelo.				
	Criterio 1A	Criterio 1B	Criterio 1C	Total
Alternativa 2. Teorema del límite central.	2	1	2	5
Alternativa 3. Organizar dataset.	3	2	3	8

Criterios para la evaluación de las ideas propuestas para la creación de la aplicación.

Nivel	1	2	3
Criterio 2A: facilidad de uso.	Baja: no se logra entender el uso del mapa en la aplicación.	Media: es difícil entender la aplicación en el uso del mapa.	Alta: se puede navegar con facilidad en la aplicación con el uso del mapa.
Criterio 2B: propiedades visuales del mapa.	El mapa no se logra ver con claridad y el nivel de zoom es malo.	El mapa se ve un poco borroso pero maneja un buen zoom.	El mapa se ve muy amigable a la vista en la aplicación con un buen nivel de zoom.

Criterio 2C: herramientas que provee.	En el mapa no se pueden añadir marcadores ni trazar rutas.	En el mapa solo se puede añadir un marcador.	En el mapa se pueden trazar rutas, añadir varios marcadores, trazar polígonos.
--	--	--	--

Evaluando con los criterios anteriores las alternativas que se mantienen, obtenemos la siguiente tabla:

Evaluación de Ideas para la creación de la aplicación				
	Criterio 2A	Criterio 2B	Criterio 2C	Total
Alternativa A. BingMapProvider	2	2	3	7
Alternativa B. GoogleMapProvider	3	3	3	9
Alternativa C. YahooMapProvider	1	2	3	6

Selección.

Según las evaluaciones anteriores, la Alternativa 2. Teorema del límite central. debe seleccionarse para calcular la probabilidad de retraso de vuelos, ya que obtuvo la puntuación más baja siendo esta la mejor de acuerdo con los criterios definidos. En segundo lugar, para seleccionar el proveedor de mapas de GMap.Net se eligió la Alternativa B. GoogleMapProvider ya que obtuvo la puntuación más alta de acuerdo con los criterios definidos.

FASE 6: DIAGRAMAS

Diagrama de clases (enlace de la imagen):

<https://drive.google.com/file/d/1n2hGOdaeLC1qNlyb4SnckUYNs7mZVwIT/view?usp=sharing>

Diagrama de objetos:

<https://drive.google.com/file/d/1OpNtgKA20PyS7glT1YPzofBnrmAjDHOi/view?usp=sharing>

SÍNTESIS REFLEXIVA:

1. Especificación de requerimientos funcionales.

Nombre	R.F# 1. Seleccionar la ciudad de origen del vuelo
Resumen	El usuario podrá, haciendo click en el mapa, elegir la ciudad de abordaje.
Input	<ul style="list-style-type: none">evento (click)
Output	Se ha establecido correctamente el valor de la ciudad de origen

Nombre	R.F# 2. Seleccionar la ciudad de destino del vuelo
Resumen	El usuario podrá, haciendo click en el mapa, elegir la ciudad de abordaje.
Input	<ul style="list-style-type: none">evento (click)
Output	Se ha establecido correctamente el valor de la ciudad de destino

Nombre	R.F# 3. Mostrar probabilidad de retraso
Resumen	El usuario podrá conocer cuál es la probabilidad de que el vuelo que seleccionó se retrase cinco minutos, 10 minutos o más de 10 minutos.
Input	<ul style="list-style-type: none">evento (click)evento (click) : seleccionar opción (5 min, 10 o >10)
Output	Se muestra un double representando la probabilidad requerida.

Nombre	R.F# 4. Mostrar probabilidad de adelanto
Resumen	El usuario podrá conocer cuál es la probabilidad de que el vuelo que seleccionó se adelante cinco minutos, 10 minutos o más de 10 minutos.
Input	<ul style="list-style-type: none"> • evento (click) • evento (click) : seleccionar opción (5 min, 10 o >10)
Output	Se muestra un double representando la probabilidad requerida.

Nombre	R.F# 5. Trazar ruta
Resumen	El usuario podrá trazar una ruta en el mapa entre varios lugares seleccionados previamente en el mapa.
Input	<ul style="list-style-type: none"> • eventos (click en los lugares)
Output	Se traza una ruta en el mapa uniendo los lugares seleccionados por el usuario.