

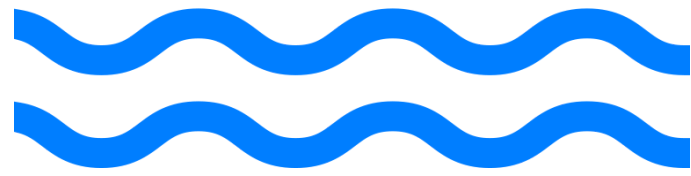
# Redegal



## Ingenierio de datos Jr

### Prueba técnica

**Fecha: 22/12/2021**



# Prueba técnica para Ingeniero de datos junior

## ¿Qué tienes que hacer?

Utilizando los datos de viajes de “Yellow Taxi” de NYC disponibles aquí<sup>1</sup>, desarrolla un programa (script o lo que tu consideres) que reciba como entrada cualquiera de los ficheros CSV mensuales disponibles, y devuelva como resultado los 10 destinos (incluyendo la información de Borough y Zone) con el mayor número de viajes que superan el percentil 0.95 en distancia de viaje para dicho mes.

El resultado (ignorando el formato, pues es totalmente libre), debería ser algo así:

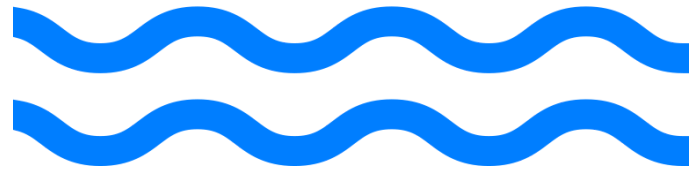
trips	end_borough	end_zone
2058	Unknown	NA
1858	Queens	JFK Airport
836	Brooklyn	Flatbush/Ditmas Park
795	Queens	LaGuardia Airport
747	Brooklyn	Crown Heights North
744	Manhattan	Clinton East
725	Brooklyn	Bay Ridge
719	Manhattan	Murray Hill
679	Brooklyn	Prospect-Lefferts Gardens
673	Manhattan	East Village

## ¿Qué tienes que enviarnos?

Un repositorio GIT (enlace a un repo de Github, Bitbucket, etc., o una copia de un repo local) con todo el código que hayas utilizado y un fichero README explicando tu aproximación al problema, si has utilizado librerías de terceros, indica las librerías que hayas utilizado (por qué y para qué), los pasos necesarios para poder ejecutar el programa, y en definitiva cualquier información que consideres relevante.

---

<sup>1</sup> <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>

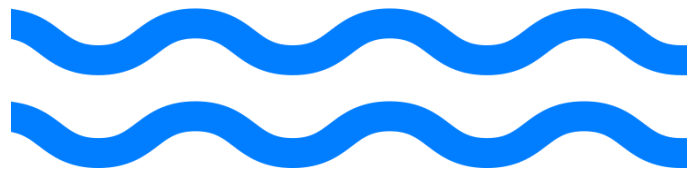


## ¿Cómo lo vamos a valorar?

1. Explicación.
2. Organización y estilo de código.
3. Facilidad de reproducción, ejecución y test.
4. Robustez de la solución.

## FAQ

- **¿Puedo utilizar cualquier lenguaje de programación?** Si. Resuelve el problema con la herramienta que te parezca más adecuada, pero recuerda que esperamos un programa o script ejecutable.
- **¿Puedo utilizar librerías de terceros?** Sí, utiliza cualquier librería que te parezca oportuna, eso sí, explícanos por qué la utilizas. Te pedimos que no utilices librerías o servicios de pago que eviten que podamos reproducir el ejemplo.
- **¿Cuánto tiempo tengo?** Tómate el tiempo que necesites, pero ten en cuenta que lo que más vamos a valorar es la explicación y aproximación utilizada. No queremos que perdáis más tiempo del necesario, y por tanto, no se trata de intentar optimizar el código al máximo ni nada por el estilo.
- **¿Qué hago si tengo más preguntas?** No dudes en escribir a Gervasio Varela en [gervasio.varela@redegal.com](mailto:gervasio.varela@redegal.com).



## A CORUÑA

Av. Linares Rivas, 56. 15005

**+34 988 549 858**

**[info@redegal.com](mailto:info@redegal.com)**