***TRABAJO INVESTIGATORIO TEORIA DEL RIESGO Y LA INCERTIDUMBRE***

*Camila Cardona*

1. **Resumen**

El objetivo primordial es evaluar el comportamiento de algunas variables al haber realizado una encuesta a 18 personas en un programa llamado R studio, dedicado a la computación de datos estadísticos, así como también gráficos de todo tipo.

1. **Palabras clave**

Programación, media, desviación estándar, probabilidades, ggplot2

1. **Introducción**

Se elaboró una encuesta con la informacion de cada uno de los miembros de la clase de teoria a la incertidumbre y riesgo, despues toda esa información obtenida acerca de la estatura, seguidores en instagram, veces que llegan tarde, y haber viso videos de mi youtuber Roccinella (experta en maquillaje) plasme toda esa informacion en excel y guarde el archivo separado por comas. En este punto Importamos el archivo separada por comas (CSV) de excel, se hizo la realizacion de comandos para poder hallar datos como la media, la desviacion estandar, los histogramas, probabilidades y graficos con paquete ggplot2.

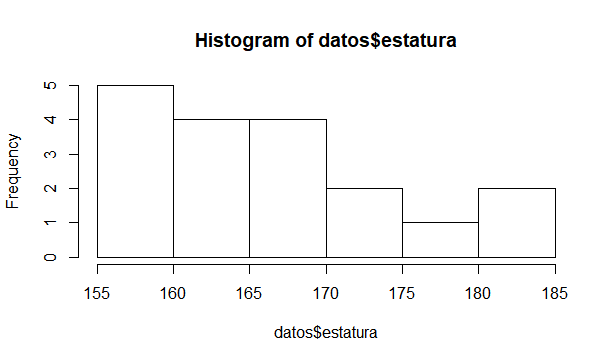
1. **Resultados**

En principio se obtienen 4 variables: Estaatura, seguidores en instagram, llegadas tarde y youtuber Roccinella, se halló la media y la desviacion estandar para aquellas variables numericas en el programa r Studio y arrojó los siguientes valores:

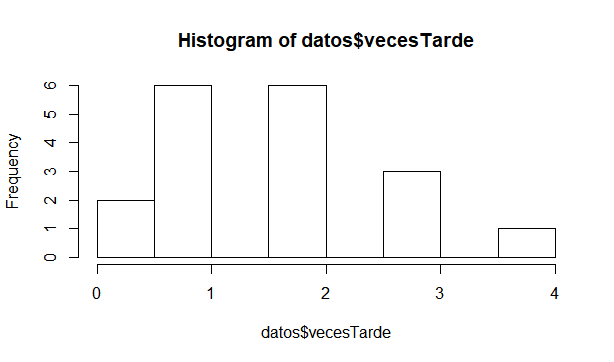


En promedio, la estatura, seguidores y llegadas tarde de los estudiantes encuestados son de 167.17, 574.61 y de 1.72 con una desviacion estandar de 9.35, 371.91, y 1. 07 respectivamente.

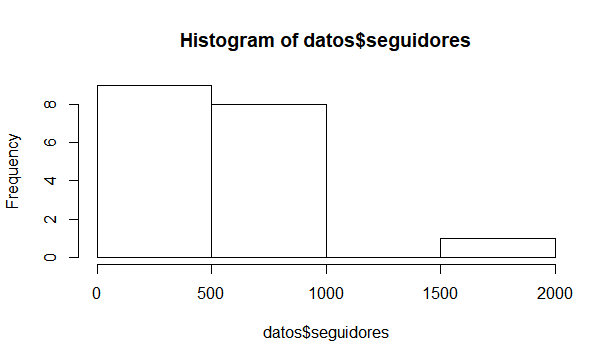
Luego se graficaron los histogramas respectivos para cada variable.



Podemos observar que la mayoria de los encuestados (5) miden entre 1,55 y 1,60. Y la minoria, solamente una persona, mide entre 1,76 y 1,80. Las demas personas estan repartidas en: 4 miden de 1,61 – 1,65. 2 de los 18 estan en un rango de 1,71 – 1,75 y otros 2 miden de 1,81 a 1,85.



En este histograma vemos que 2 personas nunca han llegado tarde a una clase y 12 personas se han retrasado por mas de 3 minutos entre 1 y 2 veces. 3 personas tan solo se han tardado 3 veces y siendo minoria, solo una persona se ha retrasado 4 veces.

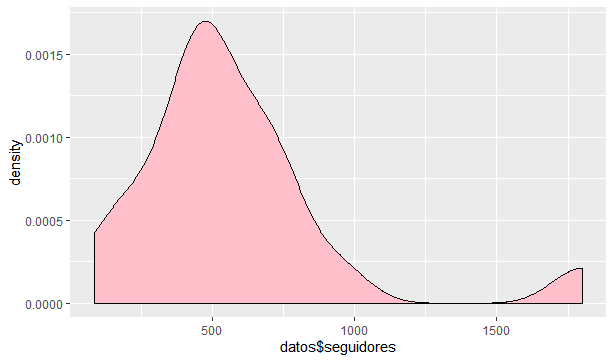
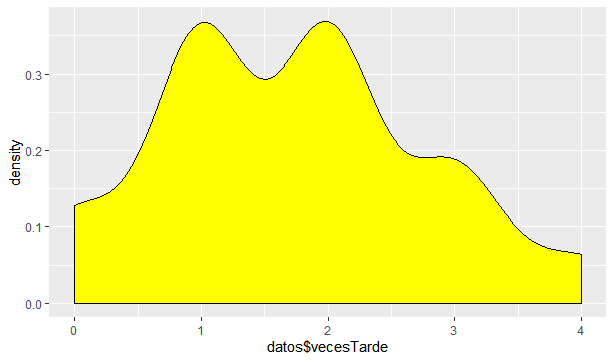
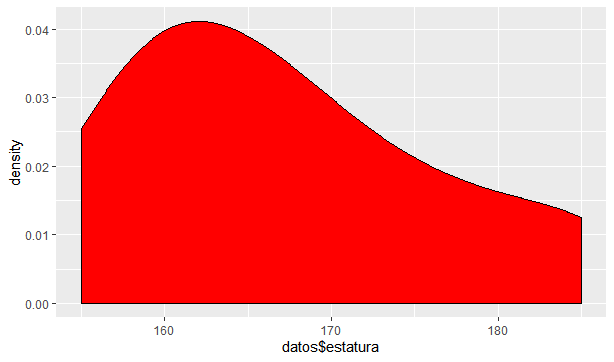


Despues de observar y estudiar detenimanete el grafico, se llega a la conclusion de que mas de 8 personas de las encuestadan tienen un rago de 0 – 500 seguidores en su instagram, 8 tienen de 501 – 1.000 y tan solo 1 persona tiene el maximo numero de seguidores que se encuentra en un intervalo de 1.501 – 2.000.

Luego se estimaron las distribuciones de probabilidad que mejor se ajustaban a las variables estudiadas.

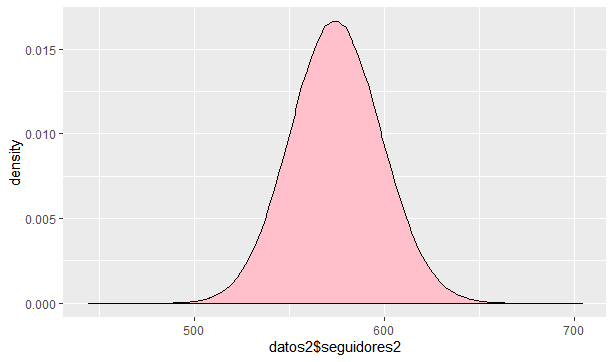
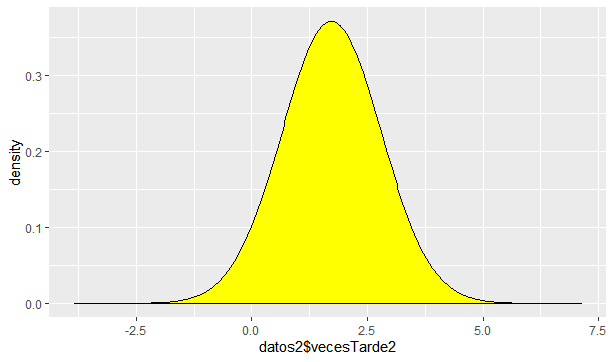
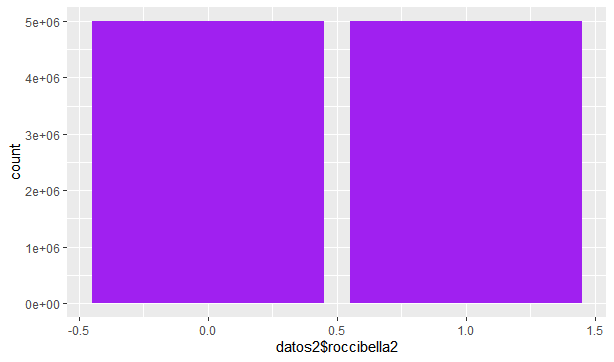
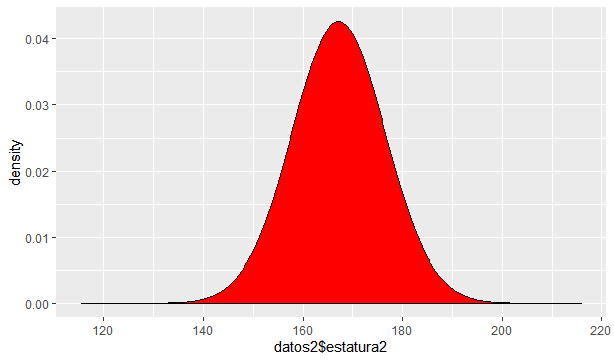
* **Estatura:** la distribucion normal ya que como es la estatura es la que mejor se acomoda a esta.
* **Seguidores:** Distribución de Poisson. La distribución de Poisson es una distribución de probabilidad discreta que se aplica a las ocurrencias de algún suceso durante un intervalo determinado. Se puede usar para calcular la llegada de pasajeros a cierto lugar.
* **Veces que llegan tarde:** sigue una distribucion normal.
* **Videos de Roccinella:** Distribucion binomial porque muestra las secuencias de n ensayos que son independientes entre si con un éxito y con un fracaso.

Luego realicé los graficos en ggplot de cada una de las variables.

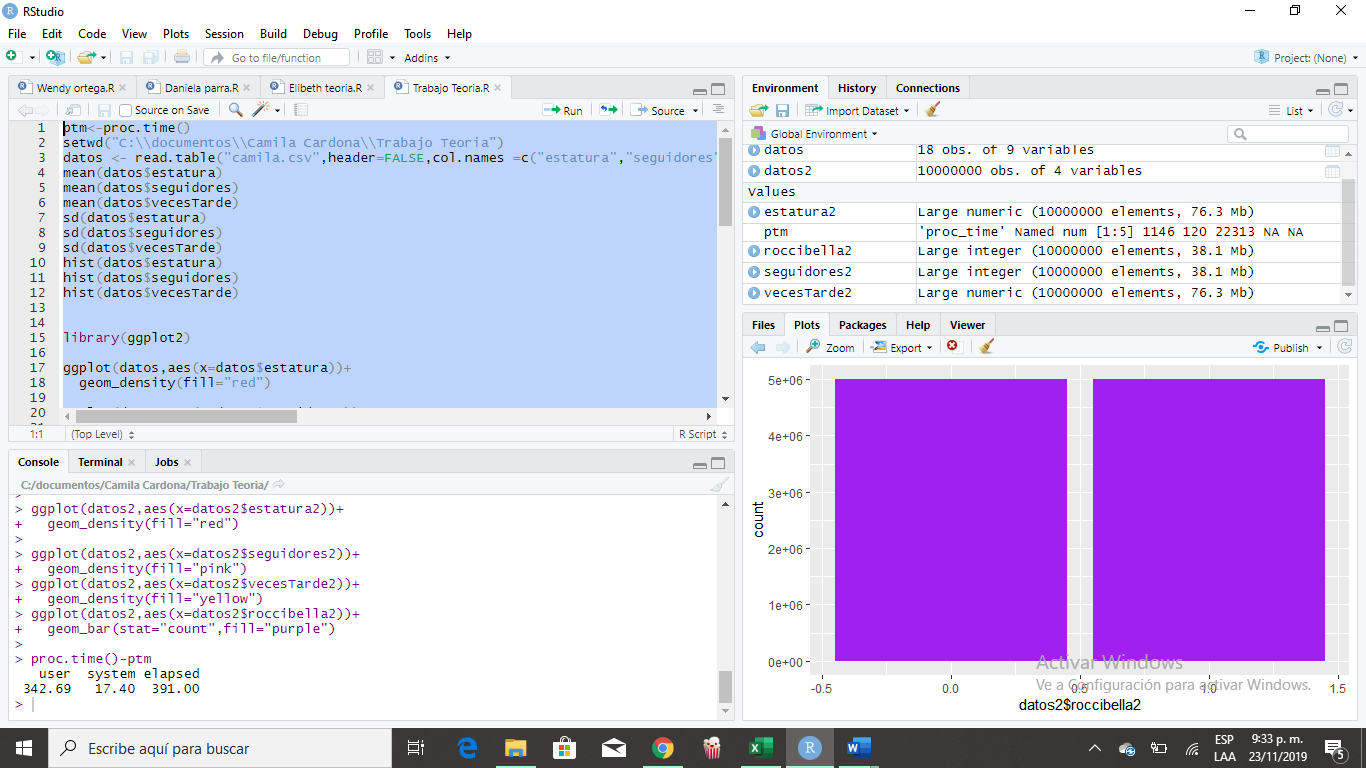


Se muestran los mismos datos que en los histogramas, pero de una manera menos basica y simple.

Luego estableci unos parametros para que los datos tuviesen un mejor ajuste, y ademas, un numero alto de simulaciones. Luego se volvio a graficar con el paquete ggplot2.



Como ultimo paso, se añade un reloj a R el cual contabiliza cuanto tarda el programa en ejecutarse. Utilice el comando proc.time. y mi resultado fue el siguiente:



1. **Conclusiones**

Con este trabajo se aprende a utilizar distintos comandos que nos facilitan mucho mas algunas situaciones en nuestra vida diara, como por ejemplo, si tenemos una tabla con alto numero de variables y datos, podemos incorporarla a nuestro R y asi hacer los calculos y graficos que necesitamos con tan solo escribiendo el comando correcto.

A modo de conclusion los graficos con el alto numero de simulaciones realizados con el paquete ggplot2 tienden a estar un poco mas hacia el medio que los realizados en primer lugar sin mencionadas simulaciones.

1. **Referencias**

* Distribucion de Poisson. <https://www.uv.es/ceaces/base/modelos%20de%20probabilidad/poisson.htm>
* Distribución binomial. <http://recursostic.educacion.es/descartes/web/materiales_didacticos/Distribucion_binomial/binomial.htm>
* Distribución normal. <http://www.matematicasvisuales.com/html/probabilidad/varaleat/normal.html>
* R studio. Reloj. <https://www.rdocumentation.org/packages/base/versions/3.6.1/topics/proc.time>