

II Jornadas Internacionales de Estadística Aplicada 5 y 6 de diciembre de 2019

VELOCIDAD DE DESCARGA DE INTERNET MOVIL

Mendieta, Daniel Alberto; Pailler, Andrea Florencia; Poma, Lautaro; Sánchez, María Emilia

Institución: Universidad Nacional de Salta – Facultad de Ingeniería

Datos de contacto: lautaropoma45@gmail.com

RESUMEN

Al inicio de la telefonía móvil los operadores usaban una tecnología que se basaba en el intercambio de datos, esto presentaba como dificultad que al ser a través de ondas aéreas se congestionaba con mucha frecuencia.

Entonces llegaron las nuevas tecnologías cuyas conexiones se realizan convirtiendo la información, en pequeños paquetes de datos, en piezas de rompecabezas que se arman al llegar a su destino.

Así, no es necesario acaparar un carril de conexión para intercambiar datos. Hoy en día, los teléfonos están siempre conectados a la red para saber si están recibiendo o no paquetes.

Como cualquier carretera también tiene horas de alto tráfico en las que muchas personas quieren conectarse a internet. Cuanto esto sucede la conexión se vuelve más lenta o incluso inexistente.

Peor aún, un estudio asegura que las "horas de más tráfico" se extienden hasta por 12 horas. Y que el 10% de los usuarios de la red móvil generan el 85% del tráfico, consumiendo sobre todo video.

Al contratar un servicio de internet para uso cotidiano acudimos a las distintas compañías que brindan internet satelital, una de las principales ventajas es su inigualable cobertura.

Se ofrecen diferentes paquetes de datos para los distintos usos que uno puede darle, pero muchas veces el servicio recibido difiere del contratado, esto se debe a distintos factores que van a ser analizados en el proyecto.

Palabras claves: velocidad de bajada, servidor, compañía, zonas.

INTRODUCCION

Los cambios en las velocidades de internet pueden deberse, entre varias cosas, a la ubicación desde donde se está usando el dispositivo, así como también puede ser debido a la compañía de la cual se trate. Muchas veces las distintas compañías de internet, de acuerdo con la zona de la ciudad de la cual se trate, varían el presupuesto que destinan a la mejora en las antenas que transmiten la señal.

A través de una aplicación se pudo obtener el mapa de la ciudad de Salta en tiempo real con todos los usuarios conectados a Internet, donde los puntos verdes marcan aquellos con una buena conexión, así como los rojos una mala conexión.

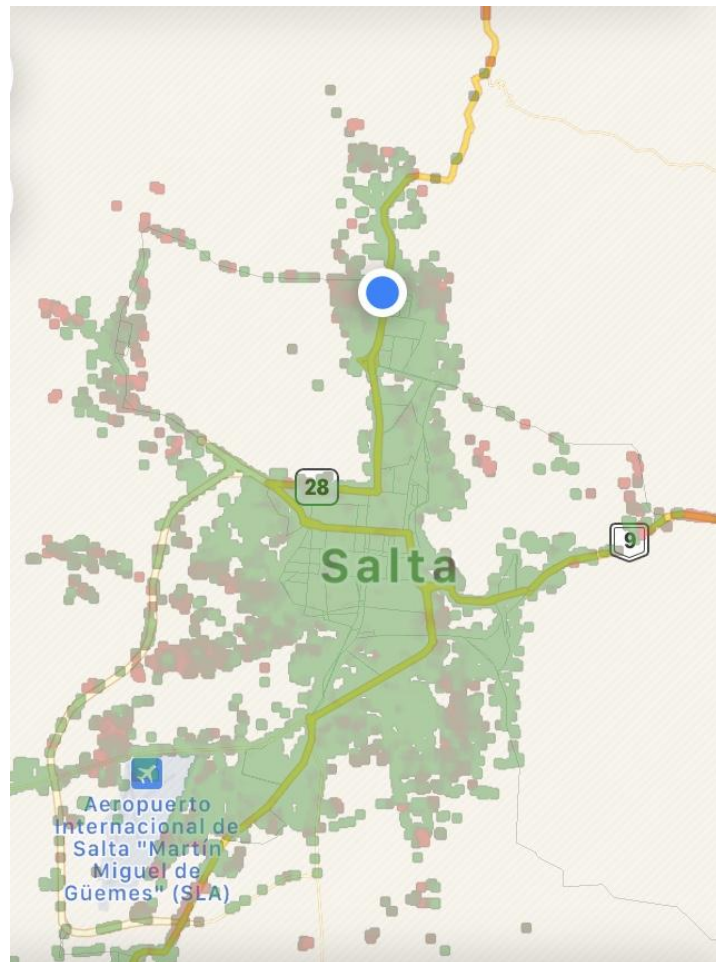


Imagen (1): Mapa de Salta que indica los usuarios conectados a Internet.

Para comenzar con el análisis se ha dividido la ciudad de Salta en cinco zonas a analizar (Norte, Sur, Centro, Este, Oeste). En un día que se consideró óptimo, en cuanto a las condiciones climatológicas, fueron tomadas en simultaneo dos medidas en cada una de las zonas de las velocidades de bajada de internet de cada una de las compañías que brindan servicio.



Universidad Nacional de Salta
FACULTAD DE INGENIERIA



FACULTAD DE
INGENIERIA
UNIVERSIDAD NACIONAL DE JUJUY



INGENIERIA
UNIVERSIDAD CATOLICA DE SALTA



Facultad de Ciencias Exactas
UNIVERSIDAD NACIONAL DE SALTA

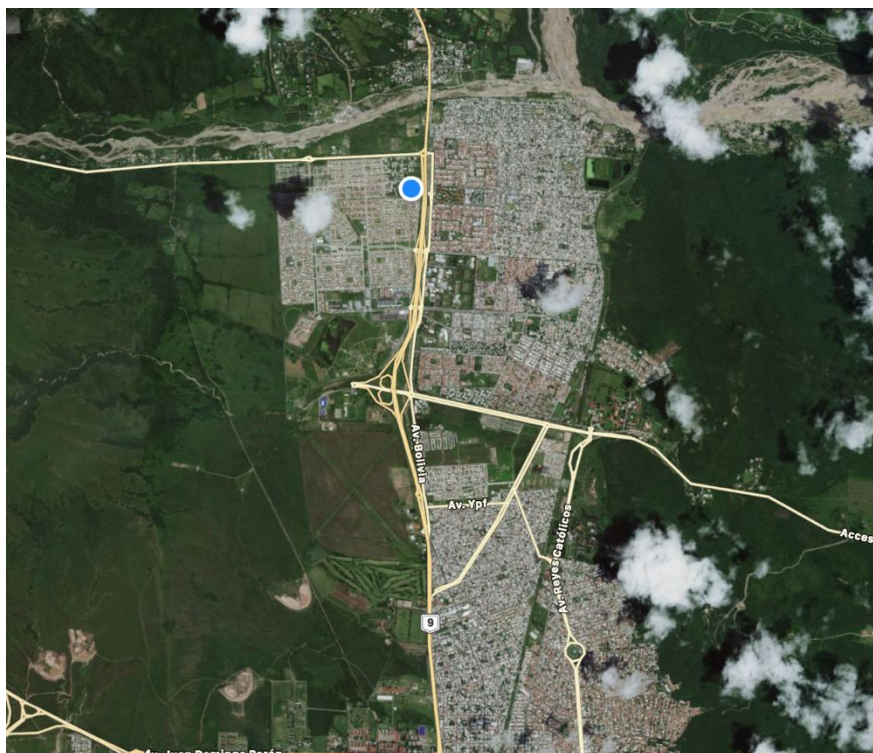


Imagen (2): Zona Norte



Universidad Nacional de Salta
FACULTAD DE INGENIERIA



FACULTAD DE
INGENIERIA
UNIVERSIDAD NACIONAL DE JUJUY



INGENIERÍA
UNIVERSIDAD CATÓLICA DE SALTA



Facultad de Ciencias Exactas
UNIVERSIDAD NACIONAL DE SALTA

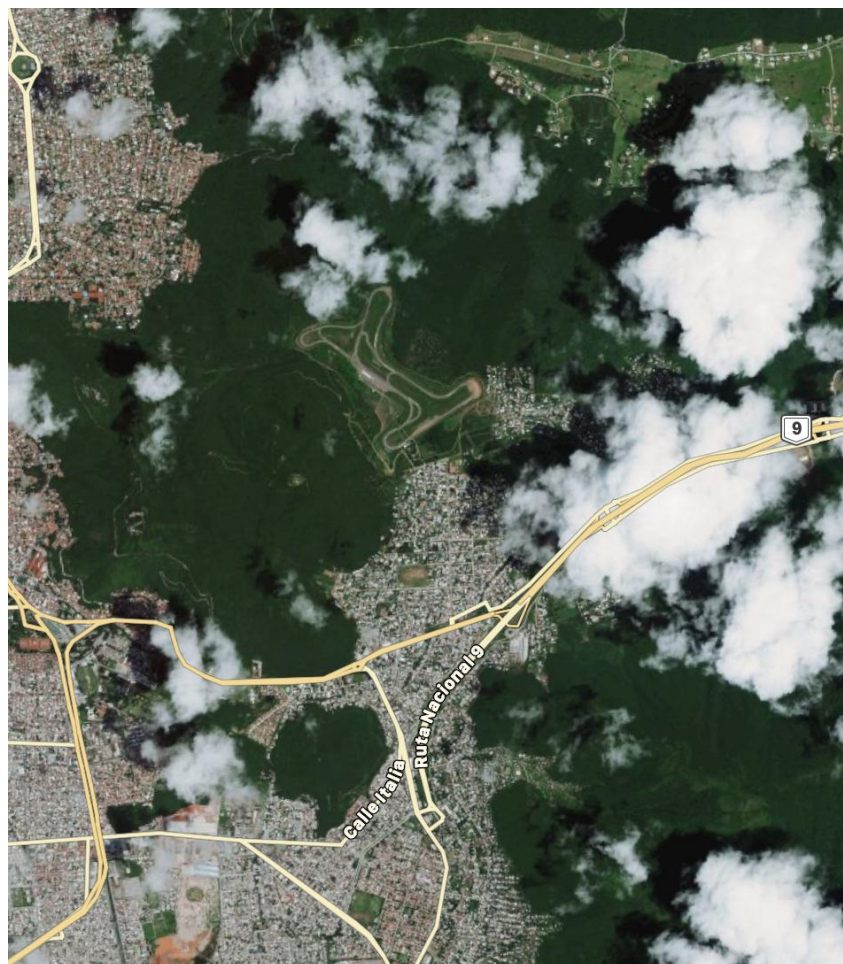


Imagen (3): Zona Este



Universidad Nacional de Salta
FACULTAD DE INGENIERIA



FACULTAD DE
INGENIERIA
UNIVERSIDAD NACIONAL DE JUJUY



INGENIERÍA
UNIVERSIDAD CATÓLICA DE SALTA



Facultad de Ciencias Exactas
UNIVERSIDAD NACIONAL DE SALTA

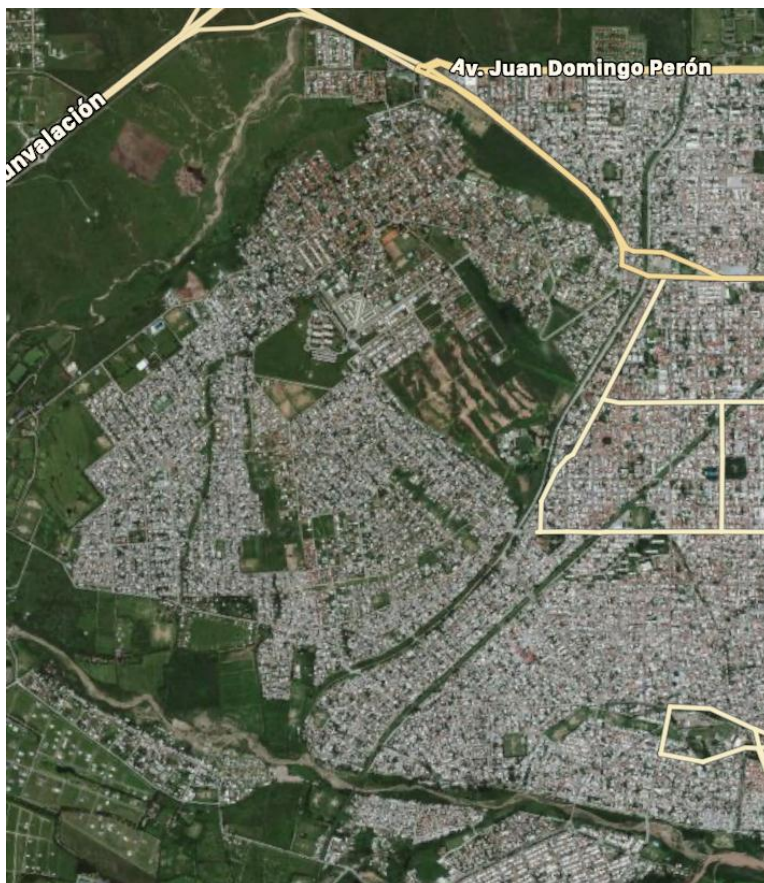


Imagen (4): Zona Oeste



Universidad Nacional de Salta
FACULTAD DE INGENIERIA



FACULTAD DE INGENIERIA
UNIVERSIDAD NACIONAL DE JUJUY



INGENIERÍA
UNIVERSIDAD CATÓLICA DE SALTA



Facultad de Ciencias Exactas
UNIVERSIDAD NACIONAL DE SALTA

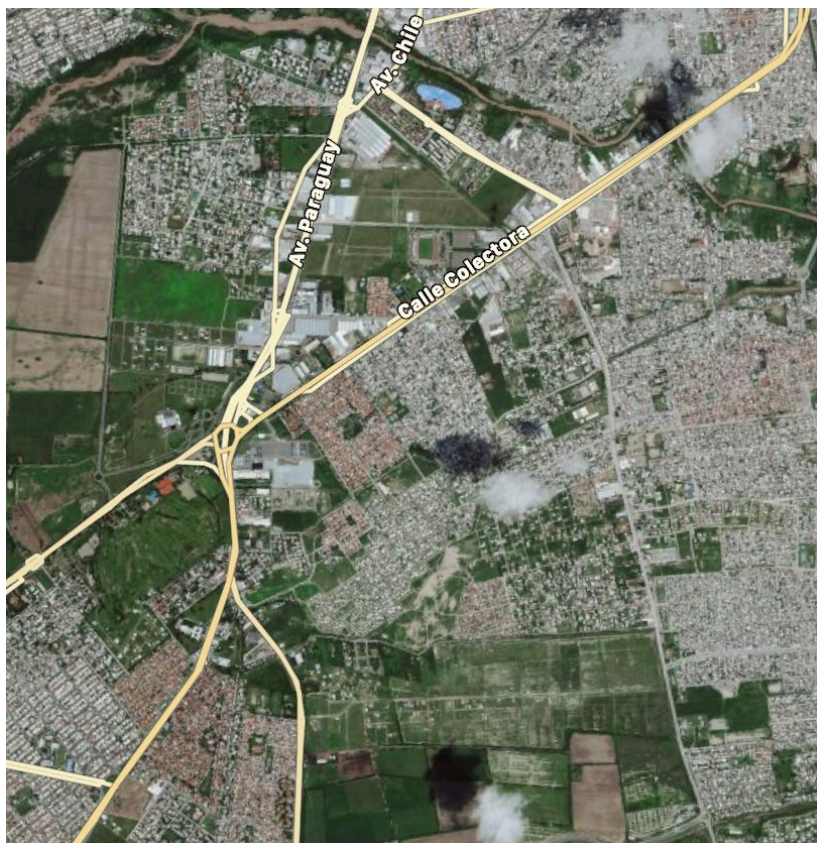


Imagen (5): Zona Sur

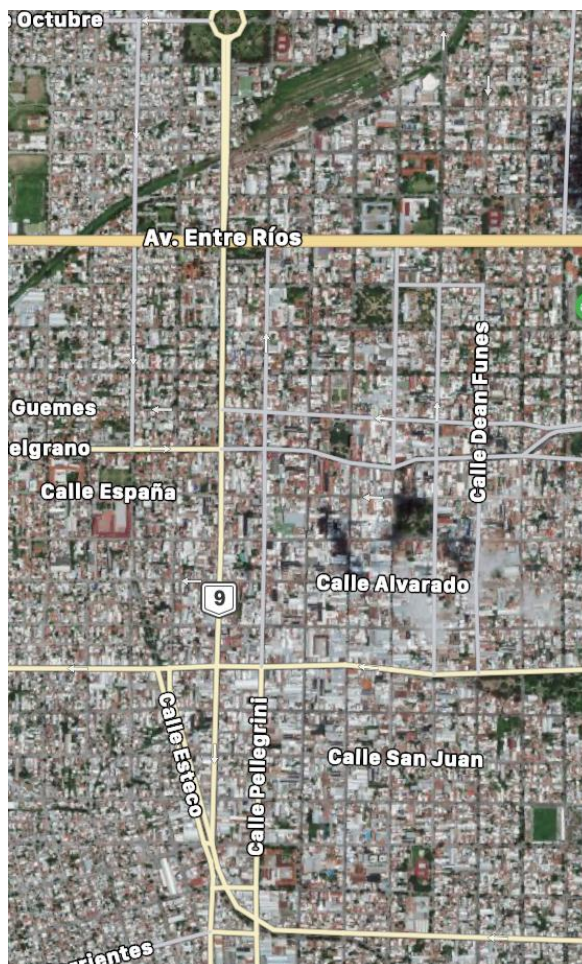


Imagen (6): Zona Centro

METODOLOGIA

Los métodos aplicados en el desarrollo de este proyecto fueron:

1. Prueba de normalidad de Shapiro Wilks

Mediante este contraste probamos que los residuales tienen una distribución normal. El procedimiento fue:

Ho: los residuales tienen una distribución normal

Ha: los residuales no tienen una distribución normal

$$w = \frac{A^2}{n * s^2}$$

donde: $A = \sum(a_{j,n} * (x_{(n-j+1)} - x_j))^2$
 $n * s^2 = \sum(x_j - \bar{x})^2$

Si $w_o < w_{crit}$, se rechaza la Ho

2. Cuadro ANOVA

Se calculan las sumas de cuadrados del error, tratamientos y total mediante:

$$SC_{\text{tratamiento}} = \frac{\sum Y_{i.}^2}{r_i} - \frac{Y_{..}^2}{n}$$

$$SC_{\text{error}} = \sum \sum Y_{ij}^2 - \sum Y_{i.}^2$$

$$SC_{\text{total}} = \sum \sum Y_{ij}^2 - \frac{Y_{..}^2}{n}$$

Siendo los grados de libertad y los cuadrados medios:

$$gL_{\text{error}} = n - t$$

$$gL_{\text{tratamiento}} = t - 1$$

$$gL_{\text{total}} = n - 1$$

$$CM_{\text{error}} = SC_{\text{error}} / gL_{\text{error}}$$

$$CM_{\text{tratamiento}} = SC_{\text{tratamiento}} / gL_{\text{tratamiento}}$$

$$CM_{\text{total}} = SC_{\text{total}} / gL_{\text{total}}$$

Y el estadístico es:

$$F_{\text{obs}} = CM_{\text{tratamiento}} / CM_{\text{error}}$$

3. Método de Tuckey

Se utiliza para la comparación de contrastes. El proceso es:

$$H_0: Y_i = Y_j ; \text{ para todo } i \neq j$$

$$H_a: Y_i \neq Y_j ; \text{ para algún par } (i, j)$$

$$\text{El número de comparaciones } T = \frac{t(t-1)}{2}$$

$$\text{Siendo el estadístico de prueba: } DHS = q_{[a; t; (n-t)]} * \sqrt{\frac{CME}{r}}$$

$$\text{El intervalo de confianza: } |Y_i - Y_j| \pm DHS$$

Los datos que fueron obtenidos de acuerdo con las distintas zonas fueron los siguientes:

	Zona norte	Zona Sur	Zona Este	Zona Oeste	Zona Centro
Compañía 1	14,55	13,67	14,56	13,28	14,14
	14,83	11,8	13,94	13,65	13,92
Compañía 2	15,47	14,88	16,67	15,13	14,95
	15,42	15,01	17,71	15,26	14,36
Compañía 3	10,57	10,74	13,24	10,05	11,2
	11,38	9,53	11,71	10,48	11,1

Tabla (1): Datos muestrales.

• Análisis de Varianza

El modelo de ANOVA utilizado es de diseño factorial, siendo los factores tomados a consideración las distintas compañías de telefonía móvil como las distintas zonas de la ciudad de Salta.

Análisis de varianza de dos factores con varias muestras por grupo						
RESUMEN	Zona norte	Zona Sur	Zona Este	Zona Oeste	Zona Centro	Total
<i>Compañía 1</i>						
Cuenta	2	2	2	2	2	10
Suma	29,38	25,47	28,5	26,93	28,06	138,34
Promedio	14,69	12,735	14,25	13,465	14,03	13,834
Varianza	0,0392	1,74845	0,1922	0,06845	0,0242	0,73876
<i>Compañía 2</i>						
Cuenta	2	2	2	2	2	10
Suma	30,89	29,89	34,38	30,39	29,31	154,86
Promedio	15,445	14,945	17,19	15,195	14,655	15,486
Varianza	0,00125	0,00845	0,5408	0,00845	0,17405	0,964382222
<i>Compañía 3</i>						
Cuenta	2	2	2	2	2	10
Suma	21,95	20,27	24,95	20,53	22,3	110
Promedio	10,975	10,135	12,475	10,265	11,15	11
Varianza	0,32805	0,73205	1,17045	0,09245	0,005	1,0336
<i>Total</i>						
Cuenta	6	6	6	6	6	
Suma	82,22	75,63	87,83	77,85	79,67	
Promedio	13,70333333	12,605	14,63833333	12,975	13,27833333	
Varianza	4,653986667	5,13515	4,917416667	5,03891	2,83665667	
ANÁLISIS DE VARIANZA						
Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Muestra	102,94952	2	51,47476	150,408376	1,1883E-10	3,682320344
Columnas	14,6696	4	3,6674	10,7160806	0,00026256	3,055568276
Interacción	4,82758	8	0,6034475	1,76326337	0,16365541	2,640796883
Dentro del grupo	5,1335	15	0,342233333			
Total	127,5802	29				

Tabla (2): Cuadro de Análisis de Varianza.

Se puede observar que no existe interacción entre los factores de compañía de internet y zona de la ciudad, lo cual habilita a continuar con el análisis.

• Comparación de a pares por Tuckey

Una vez que se observó que no había efecto de interacción, se consideró apropiado proseguir con una comparación de a pares para ver si había diferencias significativas con respecto a en qué zona se encuentra el usuario, así como también cuál es la compañía que se contrata.

Ho: Medias de las compañías son iguales								
Ha: No son iguales		DHS:	0,660080588					
Comparaciones	/yi-yj/	Limite inferior	Limite superior	Incluye al cero				
Personal - Claro	1,652	0,991919412	2,312080588	No	Existe diferencia entre las medias de descarga de las compañías evaluadas			
Personal - Movistar	2,834	2,173919412	3,494080588	No				
Claro - Movistar	4,486	3,825919412	5,146080588	No				
Ho: Medias de barrios iguales								
Ha: No son iguales		DHS:	1,20578185					
Comparaciones	/yi-yj/	Limite inferior	Limite superior	Incluye al cero				
Norte - Sur	1,09833333	-0,107448517	2,304115183	Si	Existen diferencias entre "zona este" y las demás exceptuando la norte			
Norte - Este	0,935	-0,27078185	2,14078185	Si				
Norte - Oeste	0,72833333	-0,477448517	1,934115183	Si				
Norte - Centro	0,425	-0,78078185	1,63078185	Si				
Sur - Este	2,03333333	0,827551483	3,239115183	No				
Sur - Oeste	0,37	-0,83578185	1,57578185	Si				
Sur - Centro	0,67333333	-0,532448517	1,879115183	Si				
Este - Oeste	1,66333333	0,457551483	2,869115183	No				
Este - Centro	1,36	0,15421815	2,56578185	No				
Oeste - Centro	0,30333333	-0,902448517	1,509115183	Si				

Tabla (3): Método de Tuckey.

• Intervalo de confianza para tratamientos

De Tuckey se observa que difieren significativamente entre sí las compañías, así como las zonas, por lo cual se consideró apropiado continuar con el análisis realizando intervalos de confianza para determinar cuál es la zona en la cual la velocidad de bajada de internet es de mayor nivel, así como también cuál es la compañía que brinda un mejor servicio.

Intervalo de tratamiento por compañía

Tratamiento	Media	IC del 95%	Límites
Compañía 1	13,834	$13,834 \pm 0,379$	13,45 ; 14,21
Compañía 2	15,486	$15,486 \pm 0,379$	15,11 ; 15,86
Compañía 3	11	$11 \pm 0,379$	10,62 ; 11,38
$t(0,025;27) = 2,0518$			
$S_{yi} = 0,18499$	$r=10$		

Tabla (4): Intervalos de confianza para las compañías.

De estos intervalos se puede concluir que la compañía 2 es la compañía que brinda un mejor internet de descarga siguiendo así la compañía 1 y por último la compañía 3.

Intervalo de tratamiento por zona

Tratamientos	Media	IC 95%	Límites
Zona norte	13,703	$13,703 \pm 0,23883$	13,464 ; 13,942
Zona sur	12,605	$12,605 \pm 0,23883$	12,366 ; 12,844
Zona este	14,638	$14,638 \pm 0,23883$	14,400 ; 14,677
Zona oeste	12,975	$12,975 \pm 0,23883$	12,736 ; 13,214
Zona centro	13,283	$13,283 \pm 0,23883$	13,044 ; 13,522
$t(0,025;25)=2,0505$			
$S_{yi}=0,23883$		$r=6$	

Tabla (5): Intervalos de confianza por zonas.

De este intervalo se puede concluir que la Zona Este es la que tiene mayor velocidad de descarga, siguiéndole zona norte con una diferencia poco significativa. Así también la zona con menor velocidad de la ciudad es Zona Sur.

• **Análisis de los residuales**

Se procede con el análisis de normalidad de los residuales por dos métodos: gráficamente y a través del contraste de Shapiro Wilks.

Cabe destacar que es importante realizar estas verificaciones porque al comenzar el análisis se parte de una suposición que indica que los residuales se distribuyen de manera normal con media cero y varianza constante.

En la siguiente tabla se pueden observar los valores obtenidos como los residuos.

	Modelo lineal : $e_{ij}(\text{estimado}) = y_{ij} - Y_{ij}(\text{media})$				
Residuales	-0,14	0,935	0,31	-0,185	0,11
	0,14	-0,935	-0,31	0,185	-0,11
	0,025	-0,065	-0,52	-0,065	0,295
	-0,025	0,065	0,52	0,065	-0,295
	-0,405	0,605	0,765	-0,215	0,05
	0,405	-0,605	-0,765	0,215	-0,05

Tabla (6): Residuales del modelo.

Gráficamente:

Análisis de Normalidad							
	Residuales	Distribucion Normal					Residuales
1	-0,935	0,080258221		Media Aritmetica	-1,18424E-16		-0,14
2	-0,765	0,181554134		Desviación Estándar	0,420734169		0,935
3	-0,605	0,337208893					0,31
4	-0,52	0,441776195					-0,185
5	-0,405	0,596612897					0,11
6	-0,31	0,722796944					0,14
7	-0,295	0,741563969					-0,935
8	-0,215	0,832143115					-0,31
9	-0,185	0,860832068					0,185
10	-0,14	0,897137356					-0,11
11	-0,11	0,916345388					0,025
12	-0,065	0,936956589					-0,065
13	-0,065	0,936956589					-0,52
14	-0,05	0,941532962					-0,065
15	-0,025	0,946532637					0,295
16	0,025	0,946532637					-0,025
17	0,05	0,941532962					0,065
18	0,065	0,936956589					0,52
19	0,065	0,936956589					0,065
20	0,11	0,916345388					-0,295
21	0,14	0,897137356					-0,405
22	0,185	0,860832068					0,605
23	0,215	0,832143115					0,765
24	0,295	0,741563969					-0,215
25	0,31	0,722796944					0,05
26	0,405	0,596612897					0,405
27	0,52	0,441776195					-0,605
28	0,605	0,337208893					-0,765
29	0,765	0,181554134					0,215
30	0,935	0,080258221					-0,05

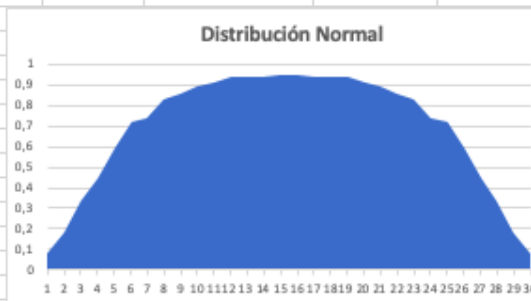


Tabla (7): Análisis de normalidad.

La grafica resultante en forma de campana indica que los residuales se distribuyen de manera normal.

Shapiro Wilks

Ho: los residuales se distribuyen normalmente.

Ha: los residuales no se distribuyen de manera normal.

f	a_{in}	$y_{(n-i+1)}$	y_i	a_y				Residuales
1	0,4254	0,935	-0,935	0,795498				-0,14
2	0,2944	0,765	-0,765	0,450432				0,935
3	0,2487	0,605	-0,605	0,300927	A	2,251067		0,31
4	0,2148	0,52	-0,52	0,223392	A ²	5,06730264		-0,185
5	0,187	0,405	-0,405	0,15147	W _{observado}	0,95420133		0,11
6	0,163	0,31	-0,31	0,10106	W _{crítico}	0,927		0,14
7	0,1415	0,295	-0,295	0,083485				-0,935
8	0,1219	0,215	-0,215	0,052417				-0,31
9	0,1036	0,185	-0,185	0,038332	W _{observado} > W _{crítico}			0,185
10	0,0862	0,14	-0,14	0,024136				-0,11
11	0,0697	0,11	-0,11	0,015334				0,025
12	0,0537	0,065	-0,065	0,006981				-0,065
13	0,0381	0,065	-0,065	0,004953				-0,52
14	0,0227	0,05	-0,05	0,00227				-0,065
15	0,0076	0,025	-0,025	0,00038				0,295
								-0,025
								0,065
								0,52
								0,065
								-0,295
								-0,405
								0,605
								0,765
								-0,215
								0,05
								0,405
								-0,605
								-0,765
								0,215
								-0,05

Como W_{observado} < W_{crítico} no se rechaza H₀, esto implica que los residuos si se distribuyen normalmente.

Tabla (8): Prueba de Shapiro Wilks.

CONCLUSIÓN

Mediante el contraste de Shapiro Wilks se corroboró el supuesto de normalidad de los errores. Es por esto que con las pruebas de hipótesis, las comparaciones de a pares y los intervalos de confianza que se pudo llegar a concluir sobre las diferencias que presentan, tanto por compañía como por zona, en las velocidades de descarga en internet móvil en la ciudad de Salta.

En el análisis de datos obtenidos se observó que las compañías no brindan la misma calidad en este servicio, el cual es un factor muy importante para los usuarios ya que las tarifas que las mismas cobran a los usuarios suelen ser elevadas y actualizadas periódicamente. Como se ha observado las empresas proveedoras no pueden brindar un servicio estable en toda una ciudad, dejando algunas de ellas bastante que desear. Es así que el usuario debe evaluar cuál de estas ofrece una cobertura que se adecue su uso.

Además, en el estudio resultó evidente que a distintas zonas de la ciudad se registran variaciones importantes lo cual puede indicar un serio problema en la calidad de las antenas de las compañías, dejando a la vista las zonas en donde se deberían invertir un mayor capital para mejorar este servicio, donde en la zona este se presenta notables diferencias en la velocidad de descarga respecto a las demás zonas, a excepción de la norte.

Como ultima observación, cabe aclarar que la calidad de los dispositivos móviles que realizan la navegación por internet es un factor influyente en este servicio. Ya que no todos tienen la misma calidad de antena receptoras y capacidad de navegación.

Bibliografía:

- Diseño de Experimentos. Robert O. Kuehl Editorial Thomson. Segunda Edición. México. 2001

Para la obtención de los datos:

- OpenSignal. <https://www.opensignal.com>
- Speedtest, test de velocidad y analisis. <https://www.testdevelocidad.es/internet-carga-descarga/>
- CoverageMap, mapeo de internet.
- Internet, articulo de Wikipedia, <https://es.wikipedia.org/wiki/Internet>