

# Ejercicio 1

# Data Science

Grupo 4  
Gulio Valenzuela  
Emily Soto  
Raul Castellanos  
Ian Jenatz

# Identificar variables

- Calificación: categórico
  - Datos de resultados sobre 40 pts
- Género: categórico
  - 1 y 2 femenino y masculino 3 IND
- Edad: continuo
- Preguntas: categórico
  - sobre 4pts las 10 preguntas

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	calificacion	genero	edad	
1		1	2	1	1	4	2	3	3	1	19	1	41
2		1	1	1	1	2	1	2	1	2	14	1	50
2		1	1	1	1	4	1	1	1	2	15	1	23
2		1	1	1	1	4	1	1	1	4	17	1	42
1		1	1	1	1	2	1	1	1	1	11	1	36
2		3	2	1	2	4	1	4	4	1	24	1	29

# Identificar datos faltantes

Especificación de  
género

Femenino masculino

Hay una edad  
que dice 999

datos que sesgan

Falta de  
contexto

no hay explicación  
de la data

Row Labels ▼	Count of genero
0	13
1	2297
2	1051
3	15
<b>Grand Total</b>	<b>3376</b>

# Datos atípicos

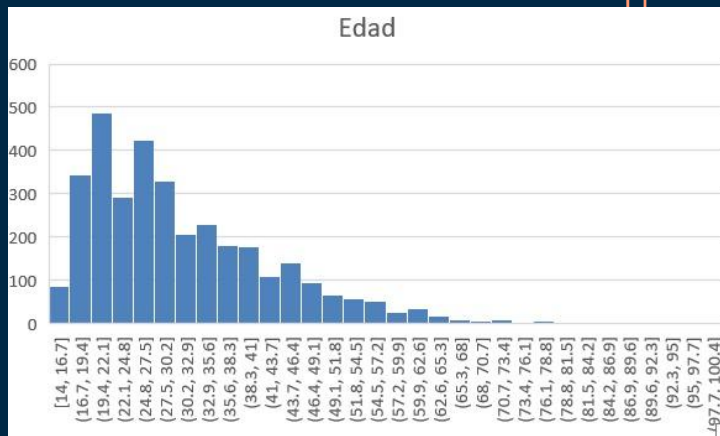
Asumiendo que  
los 0 son falta  
de respuesta

Hay una edad que dice  
números grandes y pequeños

datos que sesgan

datos atípicos de edades

de 80 para arriba



100	1
123	1
999	1
Grand Total	3376

# Análisis Univariado

Q1		Q2		Q3		Q4		Q5		Q6		Q7	
Mean	2.30628	Mean	2.225118	Mean	2.220083	Mean	1.934242	Mean	2.235782	Mean	3.085012	Mean	2.18987
Standard E	0.018828	Standard E	0.018579	Standard E	0.018787	Standard E	0.017985	Standard E	0.019319	Standard E	0.017051	Standard E	0.018575
Median	2	Median	2	Median	2	Median	2	Median	2	Median	3	Median	2
Mode	2	Mode	1	Mode	1	Mode	1	Mode	1	Mode	4	Mode	1
Standard E	1.093998	Standard E	1.079494	Standard E	1.091581	Standard E	1.044974	Standard E	1.122473	Standard E	0.990713	Standard E	1.079272
Sample Va	1.196832	Sample Va	1.165307	Sample Va	1.191549	Sample Va	1.091971	Sample Va	1.259946	Sample Va	0.981512	Sample Va	1.164828
Kurtosis	-1.17789	Kurtosis	-1.1024	Kurtosis	-1.09863	Kurtosis	-0.64634	Kurtosis	-1.21631	Kurtosis	-0.4539	Kurtosis	-1.09257
Skewness	0.246703	Skewness	0.330276	Skewness	0.340957	Skewness	0.728525	Skewness	0.340837	Skewness	-0.74512	Skewness	0.335012
Range	4	Range	4	Range	4	Range	4	Range	4	Range	4	Range	4
Minimum	0	Minimum	0	Minimum	0	Minimum	0	Minimum	0	Minimum	0	Minimum	0
Maximum	4	Maximum	4	Maximum	4	Maximum	4	Maximum	4	Maximum	4	Maximum	4
Sum	7786	Sum	7512	Sum	7495	Sum	6530	Sum	7548	Sum	10415	Sum	7393
Count	3376	Count	3376	Count	3376	Count	3376	Count	3376	Count	3376	Count	3376

Q8		Q9		Q10		calificacion		genero		edad	
Mean	2.287026	Mean	2.45468	Mean	2.509182	Mean	23.44727	Mean	1.316351	Mean	31.30332
Standard E	0.018911	Standard E	0.020191	Standard E	0.02061	Standard E	0.134914	Standard E	0.008306	Standard E	0.350346
Median	2	Median	2	Median	3	Median	23	Median	1	Median	28
Mode	1	Mode	1	Mode	4	Mode	21	Mode	1	Mode	20
Standard E	1.098804	Standard E	1.173165	Standard E	1.197493	Standard E	7.838984	Standard E	0.482628	Standard E	20.35629
Sample Va	1.207369	Sample Va	1.376316	Sample Va	1.43399	Sample Va	61.44966	Sample Va	0.23293	Sample Va	414.3785
Kurtosis	-1.18642	Kurtosis	-1.38529	Kurtosis	-1.47038	Kurtosis	-0.82276	Kurtosis	-0.74789	Kurtosis	1513.407
Skewness	0.225793	Skewness	-7.2E-05	Skewness	-0.04136	Skewness	0.173992	Skewness	0.804436	Skewness	32.06327
Range	4	Range	4	Range	4	Range	40	Range	3	Range	985
Minimum	0	Minimum	0	Minimum	0	Minimum	0	Minimum	0	Minimum	14
Maximum	4	Maximum	4	Maximum	4	Maximum	40	Maximum	3	Maximum	999
Sum	7721	Sum	8287	Sum	8471	Sum	79158	Sum	4444	Sum	105680
Count	3376	Count	3376	Count	3376	Count	3376	Count	3376	Count	3376

# Análisis Multivariado

