

Participant	Alex	Pol	Tommaso	Eric	He
1. Risk Plan and Gantt(2) D2		X			X
2. Preprocessing					
2.1 Feature selection (2) (D1)					X
2.2 Stadistical sampling (D2)					
2.3 Redundacy data (5) (D4)			X		
2.4 Technical review of data (D1)					
2.5 Transform categorical into one-hot encoding (4) (D3)		X			
2.6 Impute missing data (2) (D4)					X
2.7 Data visualization (1) (D5)	X				
2.8 Factors to describe (3) (D3)				X	
3. PCA (4)		X			
4. KNN (5)			X		
5. Document (word style)	X	X	X	X	X
6. Powerpoint	X	X	X	X	X
DURACION: de menor a mayor D1-D5					
2.1: [FEATURE SELECTION]: Seleccionar los datos que consideremos importantes. Get a DIMENSIONALITY REDUCTION.					
2.2: [STADISTICAL SAMPLING]: Si hay mucho sample (n): hay que seguir una STADISTICAL SAMPLING para coger una pequena parte de la poblacion.					
2.3 [REDUNDANCY DATA]: Para var/features. Explotatory analisis, basics statistic. Eliminar los reduntantes y representar los mas relevantes.					
2.4 [TECHNICAL REVIEW OF DATA]: revisar que los valores en cada columna son iguales.					
2.5 [T					
2.6 [Missing data]: En algun caso tenemos que encontrar algun valor para estos datos.					
2.7 plots, barplots, recharts, correlation matrix, contigency tables.		COMENTARIO: esto va relacionado con redundancy data, ya que para eliminar datos redundantes hay que analizarlos primero			
2.8 [Factors]					