

# Planificación del proyecto



**Timing:** Se dividió el proyecto en obtención de información, estudio estado del arte, obtención del data set, preprocesado, minería de datos y exploración de resultados, **outline** clásico de ciencia de datos.



**Personal:** El autor del proyecto y las directoras del mismo.



**Hardware:** Mac Pro 2,6 GHz Intel Core i5 8 GB 1600 MHz DDR3. Cluster con 4 nodos de Intel Xeon E5-2665 de 8 núcleos.



**Software:** TexShop, TextEdit, Rstudio, Spark, PySpark, PyCharm, nltk, Numbers, MongoDB.



# Metodología

