

# Instrucciones para el Proyecto

## Objetivos

El proyecto consistirá en hacer un análisis estadístico bayesiano completo para unos datos. La temática del trabajo será sobre los datos derivados del coronavirus. No obstante si alguien quiere plantear alguna temática distinta que se ponga en contacto con los profesores.

Se trata de que escojáis vosotros mismos qué objetivo u objetivos queréis responder. A modo de ejemplo podéis intentar respuesta a preguntas del tipo:

- ¿Cuáles son los casos reales de coronavirus en cada país (o regiones de un país)?
- ¿Hay diferencia en el porcentaje de población infectada por países?, ¿de qué depende?
- La tasa de mortalidad ¿en qué medida depende de variables como: la población mayor de 65 años, número médicos por habitante, número de camas de hospital por habitante, número de UCIs por habitante, etc.?
- ¿Cómo afectan las medidas de confinamiento en los diferentes países?
- ¿Cómo afectan cuándo se toman las medidas de confinamiento respecto a la mortalidad?
- ¿Cómo afectan las medidas de desconfinamiento en la aparición de nuevos casos?
- ¿Se puede estimar el efecto del coronavirus a partir de los datos de muertes por todas las causas? ¿Existen diferencias entre esta estimación y los datos de coronavirus?
- etc.

## Evaluación

El trabajo se realizará preferiblemente en **grupos de dos**. Os pediremos dos entregas, una **primera entrega el 18 de mayo que tendrá un peso del 20%** y consistirá en la descripción de los datos así como un primer análisis bayesiano inicial (el objetivo es poder hacer un seguimiento de vuestros proyectos y que tengáis un primer feedback) y una **segunda entrega el 10 de junio que tendrá un peso del 80%** y consistirá en la realización completa del proyecto.

## Posibles fuentes de datos

A continuación os mostramos algunos enlaces a bases de datos que podéis utilizar para vuestro trabajo. También podéis buscar más datos por vuestra cuenta y compartirlos en el fórum de la asignatura:

- En [esta página](https://github.com/CSSEGISandData/COVID-19) (<https://github.com/CSSEGISandData/COVID-19>) de la Johns Hopkins University podéis encontrar [datos](https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data) ([https://github.com/CSSEGISandData/COVID-19/tree/master/csse\\_covid\\_19\\_data](https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data)) relacionados con el coronavirus por países y en detalle para EEUU. También hay un dashboard interactivo.
- <https://github.com/rubenfcasal/COVID-19> En esta página hacen una recopilación de diferentes fuentes de datos de COVID-19 en España.
- [MoMo](https://momo.isciii.es/public/momo/dashboard/momo_dashboard.html#documentacion) ([https://momo.isciii.es/public/momo/dashboard/momo\\_dashboard.html#documentacion](https://momo.isciii.es/public/momo/dashboard/momo_dashboard.html#documentacion)) es un sistema para la vigilancia de la mortalidad diaria por todas las causas en España. Desde la pestaña “Datos” podéis descargar un archivo csv con información general y por comunidades autónomas.

## Contenido de las entregas

En cuanto a la extensión de las entregas no existen criterios preestablecidos. Sin embargo, es razonable suponer que la primera entrega puede tener entre 5 y 10 páginas y la segunda entrega entre 10 y 20 páginas. Además, podéis agregar Apéndices (por ejemplo, con información adicional).

El formato de los informes deberá ser pdf. En la segunda entrega se deberán incluir también los ficheros con los scripts utilizados para realizar todos los cálculos (R, stan...) y los archivos de datos.

A continuación y a modo orientativo proponemos las siguientes estructuras para cada entrega, pero en ningún caso es obligatorio ceñirse a ella:

### Primera entrega:

1. Planteamiento de un objetivo sencillo
2. Breve descriptivo de los datos utilizados
3. Un análisis Bayesiano simple (utilizando Stan o similar)
4. Presentación e interpretación de los resultados

### Segunda entrega:

#### 1. Resumen

Breve resumen del problema, análisis y conclusiones (10 líneas como máximo)

#### 2. Introducción

Descripción del problema y justificación de su elección.

Objetivos, alcance del trabajo y otra información relevante.

Variable de respuesta y variables explicativas.

Presentación de datos: una tabla con una pequeña muestra de los datos utilizados.

### **3. Análisis**

Análisis exploratorio de datos.

Análisis Bayesiano (presentación de los modelos, validación del modelo(s), interpretación del modelo final seleccionado, predicción...)

### **4. Conclusión**

Principales conclusiones en lenguaje comprensible para los no estadísticos.

Dificultades encontradas.

Posibles extensiones del análisis.

### **5. Bibliografía**