**TRABAJO PRACTICO FINAL**

Notas:

* El TP es grupal y deberá enviarse **por e-mail a** [**sperez@uno.edu.ar**](mailto:sperez@uno.edu.ar) con copia a todos los integrantes del grupo.
* Enrega: a lo sumo el **viernes 6/6 a las 23:59 hs**.
* Entreguen un documento word o pdf con sus respuestas y también el script de R utilizado en la primera parte.

Los datos del archivo **seguros.xls** corresponden a datos recopilados personas a quienes se registraron algunas características personales, la región donde vive, el índice de masa corporal, sus ingresos y el premio pagado por un seguro de salud (ambos en miles de pesos).

**Parte 1: análisis exploratorio**

1. Identificar el tipo de cada una de las variables registradas (cuali, cuanti).
2. Decidir si hay datos faltantes, en cuyo caso eliminar los casos.
3. Realizar boxplots para cada variable cuantitativa, separando según la/las variables cualitativas que considere importantes.
4. Identificar si hay datos muy atípicos, en cuyo caso eliminar los casos.
5. Hallar las medidas resumen de cada variable cuantitativa, separando según lo hecho en el ítem anterior.
6. Calcular la correlación entre todos los pares posibles de variables numéricas.

**Parte 2: inferencia estadística**

1. Realizar pruebas de comparación de medias para la variable IMC según si fuma o no.
2. Realizar pruebas de comparación de medias para la variable ingreso según género. Si considera que hay alguna hipótesis de interés a probar, según lo observado en la Parte 1, realizar el test correspondiente.
3. Realizar pruebas de comparación de medias para las variables IMC, ingreso y premio según Región.
4. Probar si puede asumirse independencia entre la procedencia (Región) y la condición de fumador de la persona.