|  |  |
| --- | --- |
|  | **Universidad de Antioquia** **Facultad de Ingeniería**  **Probabilidad e Inferencia Estadística**  **Semestre 2024 - I**  **Docente: Jorge Iván Pérez García** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Nombre** |  | **Cédula** |  |
| **Nombre** |  | **Cédula** |  |

La base de datos contiene la información de un grupo de personas que acudieron a una IPS durante la última semana por diferentes razones. La base de datos contiene las siguientes variables.

* **EPS**: Hace referencia a la Entidad Promotora de Salud a la cual se encuentra afiliado el paciente.
* **Afiliación:** Hace referencia al tipo de afiliación que tiene el paciente con su EPS.
* **Sexo:** Hace referencia al sexo del paciente.
* **Triage:** Hace referencia al sistema de ​selección y clasificación de pacientes en los servicios de urgencia.
* **Presion:** Hace referencia a la presión arterial del paciente al momento de ser atendido.
* **Azucar:** Hace referencia al azúcar en la sangre del paciente al momento de ser atendido.
* **T\_Espera:** Hace referencia al tiempo que tuvo que esperar el paciente en sala de espera hasta ser atendido.

**LEA ATENTAMENTE:**

* Diligencie el nombre y la cédula cuidadosamente.
* El trabajo puede ser realizado de forma individual o en parejas.
  + Si se realizar el trabajo de **forma individual**, debe usar la base de datos que se encuentra con su número de cédula.
  + Si se decide realizar el **trabajo en parejas**, debe usar la base de datos de los dos integrantes que se encuentra con el número de cédula de los dos, y unirlas en Excel (copiar y pegar una base debajo de la otra) o en R leyendo las dos bases de datos en R, y usando la función rbind(nombre BD1, nombre BD2)
* La solución del trabajo debe realizarse a través de un informe escrito (Word u otro editor de texto) en donde se expongan los planteamientos e interpretaciones sobre los puntos que se plantean en este trabajo, teniendo siempre el contexto de los datos.
* El método de evaluación de cada punto del trabajo estará dado por, **60%** asociado a la **solución del punto**, mientras que, el **40%** restante asociado a la **interpretación en el contexto** de la base de datos.
* La solución de las preguntas debe tener los resultados obtenidos en **R** y su debida interpretación junto a los gráficos que considere convenientes.
* Para facilitar la presentación de gráficos y salidas de **R**, puede emplear la **Herramienta de Recortes** de Windows**.**
* Una vez termine el trabajo deberá enviar el trabajo (**documento de Word, y el código de R**) al correo electrónico **jivan.perez@udea.edu.co**, con el nombre del integrante o integrantes en el asunto.

**Preguntas**

Usando la información de su base de datos responda a las siguientes preguntas. **Justifique sus respuestas gráficamente y numéricamente, junto con sus debidas interpretaciones**. ¡¡¡Recuerde que las interpretaciones son el 40% del punto!!! >:c

1. **(0.5 Puntos)** Identifique y justifique el tipo y escala de medición de cada una de las variables contenidas en su base de datos. Además, realice la coerción de las variables en **R**.
2. **(1 Punto)** Realice una **descripción inicial de las variables cuantitativas** mediante resúmenes numéricos y/o gráficos y **variables cualitativas** mediante resúmenes tabulares y/o gráficos e **interprete** los hallazgos que considere más relevantes en el contexto de los datos.
3. **(1 Punto)** Cree la variable glucemia (niveles de glucosa en la sangre), la cual está definida para 3 rangos diferentes, a saber, hipoglucemia, glucemia normal, hiperglucemia, tal que

* Hipoglucemia
* Glucemia Normal
* Hiperglucemia

y muestre si hay una relación entre los niveles de glucosa en la sangre y el tiempo mediano que debe esperar el paciente en sala de espera hasta ser atendido.

**Nota:** Recuerde que para crear una variable basada en condicionales, puede usar la función ifelse(condición, cumple, no cumple).

1. **(0.5 Puntos)** ¿Hay alguna relación entre los niveles de glucosa en la sangre y el sexo que poseen los pacientes?
2. **(1 Punto)** Filtre por alguno de los tres niveles de glucosa en la sangre creados en el punto 3 y para ese grupo, determine si el tipo de afiliación influye en los tiempos de espera para la atención de los pacientes.
3. **(1 Punto)** Tomando solo las dos EPS con mayor cantidad de registros. ¿Es posible afirmar que existe algún tipo de relación entre los tiempos de espera que pasan los pacientes en sala de espera hasta ser atendidos y la presión arterial que registran al momento de ser atendidos?