|  |  |  |
| --- | --- | --- |
| Logo1 | UNIVERSIDAD CENTRO OCCIDENTAL“LISANDRO ALVARADO”DECANATO DE CIENCIAS Y TECNOLOGÍADEPARTAMENTO DE INVESTIGACIÓN DE OPERACIONES Y ESTADISTICAESTADÍSTICA II-INFORMÁTICA Y ANÁLISIS DE SISTEMAS **PROF. NORA MONSALVE** | Logo2 |

**TERCER PARCIAL – TRABAJO PRÁCTICO (35%)**

1. **Enunciado:**

Se tienen variables relacionadas con la nutrición y salud de 315 individuos. Estos datos se hacen públicos, en la dirección**: http://lib.stat.cmu.edu/datasets/Plasma\_Retinol/,** se garantiza el la confiabilidad de la información. Su tarea será determinar si existen diferencias entre los grupos de individuos estudiados. Qué recomendaciones de salud puede dar a mujeres y a hombres. Qué riesgos de salud ha encontrado después de hacer los análisis. Los datos son de un estudio transversal para investigar la relación entre las características personales y los factores dietéticos, y las concentraciones plasmáticas de retinol, betacaroteno y otros carotenoides. Los participantes del estudio fueron pacientes que se sometieron a una intervención quirúrgica electiva durante un período de tres años para biopsiar o extirpar una lesión de pulmón, colon, mama, piel, ovario o útero que no resultó cancerosa.

La base de datos llamada NutritionStudyestácontenida en el paquete **Lock5Data.** Usted deberá realizar un estudio estadístico en dos (2) partes, en una primera parte, deberá realizar un análisis descriptivo y gráfico de las variables que considere y con el cual demuestre comportamientos de centralización, variabilidad, posición y curtosis de las variables.

En una segunda parte, deberá hacer inferencia estadística usando un nivel de significancia (alpha) de 5%. Para la inferencia deberá definir lo siguiente:

* Las variables que considerará y su escala de medición.
* La verificación de la normalidad cuando corresponda (tanto gráficamente como usando el test de Shapiro, u otros test que correspondan).
* Las hipótesis a contrastar (bilateral o unilateral) en cada caso.
* Parámetros a estimar.
* Estimadores y sus modelos de probabilidad.
* Si hay dos poblaciones que se comparan en promedio, deberá comprobar si las varianzas poblaciones se asumen iguales o no.
* El criterio para decidir cuál hipótesis se elige, deberá ser a través de 2 criterios: el p-value (valor-p) y el intervalo de confianza.
* Si hay más de dos (2) poblaciones a comparar en promedio, deberá proponer la mejor forma de comparar las medias. Proponga un método para comparar más de dos (2) poblaciones.
* Si hay alguna de las variables a estudiar que no cumpla la Normalidad y deba incluirse en el estudio, de qué forma usted la incluiría para que no viole la condición de normalidad.

1. **Condiciones del trabajo:**

• El informe se recomienda no sea muy largo, se sugiere no más de 20 páginas.

• Deberá dar las recomendaciones y la presentación de los resultados en un informe técnico escrito cuyo contenido sugerido es el siguiente: resumen, introducción al tema con revisión de antecedentes, metodología estadística a seguir, conclusiones, recomendaciones y referencias bibliográficas.

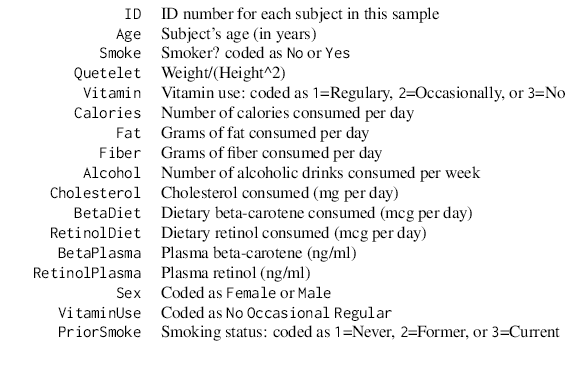
• El informe debe ser entregado en formato .doc

• El trabajo se hará en grupo de 2 ó 3 estudiantes y cada estudiante deberá manejar todo el contenido desarrollado.

• Deberá hacerse una presentación del trabajo en un tiempo aproximado de 10 minutos por estudiante, donde presentará los resultados principales, conclusiones y recomendaciones si las hay.

• Trabajos con igual o similar contenido serán penalizados.

•Se sugiere usar el software R y/o RStudio o cualquier software estadístico que aplique.

1. **Estructura de la Base de datos que deberán analizar.**

**4. Condiciones de entrega y defensa del proyecto:**

* Entrega de trabajo y defensa (una breve exposición): miércoles, 28 de Mayo 2025 en horas de clase.
* No hay prórrogas para la entrega y defensa.
* Prueba sustitutiva: Viernes, 30 de Mayo 2025.

1. **Aspectos a ser evaluados:**

|  |  |  |
| --- | --- | --- |
| **Aspectos a evaluar** | **Puntuación** | **Descripción** |
| Informe escrito | 10% | La estructura del informe escrito debe contener: resumen, antecedentes del problema, metodología estadística empleada y su sustentación teórica, presentación de resultados, discusión de los resultados, recomendaciones, referencias bibliográficas. |
| Originalidad en presentación | 5% | Estilo de redacción, de gráficos y tablas. Forma de organización de los resultados, coherencia en la discusión de los resultados. |
| Metodología Estadística empleada | 15% | Conceptualización de los métodos estadísticos, sustentación de su uso y su interpretación en la inferencia. |
| Presentación y defensa de resultados mediante una exposición | 5% | Forma de presentación de los resultados, discusión y validación de los resultados. |
| **Total evaluado** | 35% | Nota final |