|  |  |
| --- | --- |
|  | **Maestría en Ingeniería de Información (MINE)**  **Ingeniería de Sistemas y Computación**  **Curso: MINE 4101-Ciencia de datos aplicada**  **Parcial**  **Periodo: 2020-20** |

“Al entregar la solución de este examen, yo, **Carlos Andrés Salazar Aroca** con código 201816839 me comprometo a no conversar durante el desarrollo de este examen con ninguna persona diferente a los profesores del curso sobre aspectos relacionados con el examen; tampoco utilizaré algún medio de comunicación por voz, texto o intercambio de archivos, para consultar o compartir con otros, información sobre el tema del examen antes de entregarlo. Soy consciente y acepto las consecuencias que acarreará para mi desempeño académico cometer fraude en este examen”.

**Reglas**

1. El examen es estrictamente individual y tiene 100 minutos para resolverlo.
2. Utilice procedimientos explícitos y justifique su respuesta.
3. Este examen deber ser entregado por SICUAPLUS en el link destinado para tal fin, hasta la fecha máxima establecida.
4. Se debe entregar un archivo en formato html y este word

**Parcial**

**Descripción del caso:** La base de datos[[1]](#footnote-1) a utilizar contiene información sobre las exportaciones realizadas en Colombia para el mes de Diciembre de un determinado año. Los datos incluyen el lugar de origen y destino de la exportación, el medio de transporte utilizado, el peso, la cantidad y el valor de la mercancía (incluidos valores de impuestos). Al final de este enunciado se presenta la descripción detallada de la fuente. Esta fuente de es la base para resolver las preguntas del parcial que se presentan a continuación.

**I. [15%] Entendimiento del negocio y de los datos**

A partir de la exploración inicial de los datos:

1. **[5%]** Describa el resultado del perfilamiento de datos e indique el tipo de cada variable.
2. **[10%]** Describa el entendimiento que tiene del negocio a partir de los datos suministrados.

**II. [75%] Análisis y propuestas accionables**

1. **[15%]** Plantee 4 hipótesis de interés para el negocio que al resolverlas generen un beneficio. Justifique su respuesta en términos de los datos, los análisis realizados para definirlas y la organización beneficiaria de la solución.
2. **[10%]** Describa e ilustre los problemas de calidad de los datos que pueden afectar los análisis propuestos en las hipótesis.
3. **[25%]** Realice los ajustes que considere necesarios sobre los datos para realizar los análisis propuestos y justifique sus decisiones.
4. **[25%]** Valide 2 de las hipótesis planteadas en el punto anterior. Una de las hipótesis debe ser entre variables numéricas y la otra entre variables categóricas. Indique los sesgos que pueden existir con el análisis realizado para validar las hipótesis seleccionadas.

**III. [10%] Impacto en la ley de protección de datos**

1. **[10%]** Se desea generar una estrategia que permita aumentar las exportaciones de Colombia en distintos sectores de la economía Colombiana priorizando los lugares que presentan un menor número de exportaciones. Sin embargo, antes de poner en marcha cualquier estrategia, se quiere resolver la siguiente duda: ¿Qué tipo de dificultades podrían tener las iniciativas planteadas por el gobierno, en el marco de la ley de protección de datos personales colombiano?

**Descripción de variables**

|  |  |
| --- | --- |
| **Variable** | **Descripción** |
| ID | ID de la exportación |
| [COD\_SAL](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V134) | Lugar de donde sale la mercancía.  AXM ARMENIA  BAQ BARRANQUILLA  BOG BOGOTA  BGA BUCARAMANGA  CLO CALI  CTG CARTAGENA  CUC CUCUTA  MZL MANIZALES  MDE MEDELLIN  PEI PEREIRA  SMR SANTA MARTA  RCH RIOHACHA  ADZ SAN ANDRES  AUC ARAUCA  BUN BUENAVENTURA  CAR CARTAGO  IPI IPIALES  LET LETICIA  MAI MAICAO  TCO TUMACO  TRB URABA  PCR PUERTO CARRENO  INI INIRIDA  YOP YOPAL  PUU PUERTO ASIS  VUP VALLEDUPAR.  PA\_ PAMPLONA |
| VIA | Vía de transporte  1 MARITIMO 2 FERREO 3 TERRESTRE 4 AEREO 5 CORREO 6 MULTIMODAL 7 INSTALACIONES DE TRANSPORTE FIJAS (TUBERIAS, CABLE, ETC.) 8 VIAS NAVEGABLES INTERIORES 9 OTRO MODO DE TRANSPORTE |
| [CANTI](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V147) | Número de unidades del producto exportado |
| [PBK](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V148) | Peso Bruto de la mercancía |
| [PNK](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V149) | Peso neto de la mercancía |
| [FOBDOL](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V150) | Valor total de la mercancía (en miles de dólares) |
| [FOBPES](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V151) | Valor total de la mercancía (en millones de pesos) |
| [FLETES](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V153) | Valor fletes de la mercancía |
| [SEGURO](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V154) | Valor asociado a seguros de la mercancía |
| [OTROSG](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V155) | Otros gastos de la mercancía |
| PAIS | País de destino de la mercancía exportada  249 USA  239 ECUADOR  589 PERU  580 PANAMA  196 COSTA RICA  493 MEXICO  317 GUATEMALA  211 CHILE  647 REP DOMINICANA  97 BOLIVIA  149 CANADA  245 ESPANA  242 EL SALVADOR  611 PUERTO RICO  573 PAISES BAJOS  200 CURAZAO |
| **Variable** | **Descripción** |
| ID | ID de la exportación |
| [COD\_SAL](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V134) | Lugar de donde sale la mercancía.  AXM ARMENIA  BAQ BARRANQUILLA  BOG BOGOTA  BGA BUCARAMANGA  CLO CALI  CTG CARTAGENA  CUC CUCUTA  MZL MANIZALES  MDE MEDELLIN  PEI PEREIRA  SMR SANTA MARTA  RCH RIOHACHA  ADZ SAN ANDRES  AUC ARAUCA  BUN BUENAVENTURA  CAR CARTAGO  IPI IPIALES  LET LETICIA  MAI MAICAO  TCO TUMACO  TRB URABA  PCR PUERTO CARRENO  INI INIRIDA  YOP YOPAL  PUU PUERTO ASIS  VUP VALLEDUPAR.  PA\_ PAMPLONA |
| VIA | Vía de transporte  1 MARITIMO 2 FERREO 3 TERRESTRE 4 AEREO 5 CORREO 6 MULTIMODAL 7 INSTALACIONES DE TRANSPORTE FIJAS (TUBERIAS, CABLE, ETC.) 8 VIAS NAVEGABLES INTERIORES 9 OTRO MODO DE TRANSPORTE |
| [CANTI](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V147) | Número de unidades del producto exportado |
| [PBK](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V148) | Total Kilos Bruto de la mercancía |
| [PNK](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V149) | Peso neto de la mercancía |
| [FOBDOL](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V150) | Valor total de la mercancía (en dólares) |
| [FOBPES](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V151) | Valor total de la mercancía (en pesos) |
| [FLETES](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V153) | Valor fletes de la mercancía |
| [SEGURO](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V154) | Valor asociado a seguros de la mercancía |
| [OTROSG](http://microdatos.dane.gov.co/index.php/catalog/472/datafile/F3/V155) | Otros gastos de la mercancía |
| PAIS | País de destino de la mercancía exportada  249 USA  239 ECUADOR  589 PERU  580 PANAMA  196 COSTA RICA  493 MEXICO  317 GUATEMALA  211 CHILE  647 REP DOMINICANA  97 BOLIVIA  149 CANADA  245 ESPANA  242 EL SALVADOR  611 PUERTO RICO  573 PAISES BAJOS  200 CURAZAO |

1. Los datos fueron tomados de la página del DANE (Colombia-Estadística de Exportaciones-EXPO-2011 a 2017) [↑](#footnote-ref-1)