

# CONTROL #2

## Python para el análisis de datos

Fecha publicación: VIERNES 14 de JUNIO de 2024

Fecha entrega: **JUEVES 27 DE JUNIO DE 2024**

Entregas posteriores a la fecha de entrega tienen un descuento 10 décimas por día de atraso.

**📖 Importante. Lea detenidamente el enunciado. Le solicitamos encarecidamente que lea las instrucciones antes de comenzar a resolver los problemas. La fecha de entrega es el JUEVES 27 de JUNIO hasta las 23:59hrs.**

### INSTRUCCIONES

1. Resuelva **los dos problemas planteados en el control #2**. Una vez resuelto, debe subir **cada archivo con su respectivo nombre (ej. problema\_1.py, problema\_2.py)** a la página de webcursos con el link: "CONTROL\_ENTREGA\_2"
2. Este control tiene un puntaje máximo de 60 puntos. Cada pregunta tiene 30 puntos
3. Este control debe ser resuelto en forma individual.
4. Asegúrese de leer todos los pasos anteriores. Recuerde que debe resolver esta tarea en Python (solo en Python)
5. Revise nuevamente desde el primer punto de este instructivo si algo no quedó claro.

## PROBLEMA 1 (30 PUNTOS)

Usted se ha incorporado recientemente como un científico de datos a una empresa de exportaciones de frutas y vegetales, con ventas de preferencia al mercado Europeo, Oceanía y Estados Unidos. La empresa lleva varios años en el mercado, pero desconoce cuáles son los mercados con mayor volumen de venta, y la preferencia de un producto por país. Un ejemplo de la estructura de datos se muestra en la Fig.2.

|    | A        | B        | C          | D       | E        | F              |
|----|----------|----------|------------|---------|----------|----------------|
| 1  | Order ID | Product  | Category   | Amount  | Date     | Country        |
| 2  | 1        | Carrots  | Vegetables | \$4'270 | 06-01-16 | United States  |
| 3  | 2        | Broccoli | Vegetables | \$8'239 | 07-01-16 | United Kingdom |
| 4  | 3        | Banana   | Fruit      | \$617   | 08-01-16 | United States  |
| 5  | 4        | Banana   | Fruit      | \$8'384 | 10-01-16 | Canada         |
| 6  | 5        | Beans    | Vegetables | \$2'626 | 10-01-16 | Germany        |
| 7  | 6        | Orange   | Fruit      | \$3'610 | 11-01-16 | United States  |
| 8  | 7        | Broccoli | Vegetables | \$9'062 | 11-01-16 | Australia      |
| 9  | 8        | Banana   | Fruit      | \$6'906 | 16-01-16 | New Zealand    |
| 10 | 9        | Apple    | Fruit      | \$2'417 | 16-01-16 | France         |
| 11 | 10       | Apple    | Fruit      | \$7'431 | 16-01-16 | Canada         |
| 12 | 11       | Banana   | Fruit      | \$8'250 | 16-01-16 | Germany        |
| 13 | 12       | Broccoli | Vegetables | \$7'012 | 18-01-16 | United States  |
| 14 | 13       | Carrots  | Vegetables | \$1'903 | 20-01-16 | Germany        |
| 15 | 14       | Broccoli | Vegetables | \$2'824 | 22-01-16 | Canada         |
| 16 | 15       | Apple    | Fruit      | \$6'946 | 24-01-16 | France         |
| 17 | 16       | Banana   | Fruit      | \$2'320 | 27-01-16 | United Kingdom |
| 18 | 17       | Banana   | Fruit      | \$2'116 | 28-01-16 | United States  |
| 19 | 18       | Banana   | Fruit      | \$1'135 | 30-01-16 | United Kingdom |
| 20 | 19       | Broccoli | Vegetables | \$3'595 | 30-01-16 | United Kingdom |
| 21 | 20       | Apple    | Fruit      | \$1'161 | 02-02-16 | United States  |
| 22 | 21       | Orange   | Fruit      | \$2'256 | 04-02-16 | France         |
| 23 | 22       | Banana   | Fruit      | \$1'004 | 11-02-16 | New Zealand    |
| 24 | 23       | Banana   | Fruit      | \$3'642 | 14-02-16 | Canada         |

**Figura 1.** Ejemplo de planilla de datos con exportaciones de frutas y vegetales

Dados sus conocimientos en Python (y Pandas), se le ha solicitado conocer la siguiente información.

1. [5 pts] Dentro de la categoría de frutas, indique el número frutas vendidas según el tipo de frutas por país.
2. [5 pts] ¿Cuál es la fruta con mayor ventas (Amount) en total en todos los países.
3. [10 pts] ¿Cuál es el monto promedio de ventas de cada fruta?
4. [10 pts] ¿Cuál es el país que compra la mayor cantidad de Zanahorias?

## PROBLEMA 2 (30 PUNTOS)

En los últimos años el número de deportistas profesionales se ha incrementado. Esto ha significado que tanto la estatura como el peso de los atletas se ha ido moldeando según la disciplina deportiva. Empleando la información expuesta en el archivo “athlete\_events.csv”, resuelva las siguientes preguntas:

| ID | Name                     | Sex | Age | Height | Weight | Team           | NOC | Games       | Year | Season | City        | Sport         |
|----|--------------------------|-----|-----|--------|--------|----------------|-----|-------------|------|--------|-------------|---------------|
| 1  | A Dijiang                | M   | 24  | 180    | 80     | China          | CHN | 1992 Summer | 1992 | Summer | Barcelona   | Basketball    |
| 2  | A Lamusi                 | M   | 23  | 170    | 60     | China          | CHN | 2012 Summer | 2012 | Summer | London      | Judo          |
| 3  | Gunnar Nielsen Aaby      | M   | 24  | NA     | NA     | Denmark        | DEN | 1920 Summer | 1920 | Summer | Antwerpen   | Football      |
| 4  | Edgar Lindenau Aabye     | M   | 34  | NA     | NA     | Denmark/Sweden | DEN | 1900 Summer | 1900 | Summer | Paris       | Tug-Of-War    |
| 5  | Christine Jacoba Aaftink | F   | 21  | 185    | 82     | Netherlands    | NED | 1988 Winter | 1988 | Winter | Calgary     | Speed Skating |
| 5  | Christine Jacoba Aaftink | F   | 21  | 185    | 82     | Netherlands    | NED | 1988 Winter | 1988 | Winter | Calgary     | Speed Skating |
| 5  | Christine Jacoba Aaftink | F   | 25  | 185    | 82     | Netherlands    | NED | 1992 Winter | 1992 | Winter | Albertville | Speed Skating |
| 5  | Christine Jacoba Aaftink | F   | 25  | 185    | 82     | Netherlands    | NED | 1992 Winter | 1992 | Winter | Albertville | Speed Skating |
| 5  | Christine Jacoba Aaftink | F   | 27  | 185    | 82     | Netherlands    | NED | 1994 Winter | 1994 | Winter | Lillehammer | Speed Skating |
| 5  | Christine Jacoba Aaftink | F   | 27  | 185    | 82     | Netherlands    | NED | 1994 Winter | 1994 | Winter | Lillehammer | Speed Skating |
| 6  | Per Knut Aaland          | M   | 31  | 188    | 75     | United States  | USA | 1992 Winter | 1992 | Winter | Albertville | Cross Country |
| 6  | Per Knut Aaland          | M   | 31  | 188    | 75     | United States  | USA | 1992 Winter | 1992 | Winter | Albertville | Cross Country |
| 6  | Per Knut Aaland          | M   | 31  | 188    | 75     | United States  | USA | 1992 Winter | 1992 | Winter | Albertville | Cross Country |
| 6  | Per Knut Aaland          | M   | 31  | 188    | 75     | United States  | USA | 1992 Winter | 1992 | Winter | Albertville | Cross Country |
| 6  | Per Knut Aaland          | M   | 33  | 188    | 75     | United States  | USA | 1994 Winter | 1994 | Winter | Lillehammer | Cross Country |
| 6  | Per Knut Aaland          | M   | 33  | 188    | 75     | United States  | USA | 1994 Winter | 1994 | Winter | Lillehammer | Cross Country |
| 6  | Per Knut Aaland          | M   | 33  | 188    | 75     | United States  | USA | 1994 Winter | 1994 | Winter | Lillehammer | Cross Country |

Figura 1. Ejemplo de planilla de datos del archivo **athlete\_events.csv**

1. [10pts] Grafique la relación peso/tamaño de cada disciplina (agrupe los datos por disciplina).
2. [10pts] Determine cuál es la disciplina que tiene la mayor relación altura/peso del conjunto de disciplinas del problema anterior.
3. [10pts] Grafique por año en la dispersión altura/peso de la especialidad de Ciclismo.