SkilloVilla

Assignment
Inferential Statistics

SkilloVilla

BlueBull Marketing

BlueBull is a health drink manufacturing company and BlueBull has observed that in some kids who consume BlueBull, there is an increase in height of kids compared to the kids who don't consume BlueBull. And now the CEO of BlueBull wants to market this product as a drink that can cause an increase in the height of children compared to the regular population who don't consume it. Calling it "The height enhancer".

Can you help the CEO, before marketing BlueBull as a "height enhancer", to ascertain if consuming BlueBull gives you an advantage in increasing the height of children. Now, why do we need to ascertain the truth before marketing the product? Because a false marketing narrative can lead to the damage of brand name and government regulators would fine the company for false advertising.

Now an executive from BlueBull has set up an experiment with 2000 kids without any pre-medical conditions, with 1000 kids in the Test group that consumes BlueBull and 1000 kids in the control group who do not consume any health drink including BlueBull.

The kids in the experiment have been monitored for 12 months with height measurements taken at the end of each month.

Now how do we help the CEO make a decision on marketing BlueBull as "The height enhancer" with this data?

- Can the CEO market this product across India that BlueBull is a height enhancer ?
- In which states of India can the marketing be done by BlueBull based on the experiment?
- Can the CEO market this product across gender and age groups?
- 12 months is really a long time to conclude an experiment and is there a faster way to conclude the experiment for this marketing problem?
- In what categories could we absolutely say that BlueBull works as a height enhancer and in what other categories can we say that BlueBull doesn't as a height enhancer and in what categories do we need to investigate further.?

SkilloVilla

Dataset:

Dataset has two sheets: Test & Control. Both the sheets have 1000 rows of data and each have the same 20 columns.

Columns in Data:

Kid Name: Name of the child participating in the experiment **Age(In Years):** Age of the child in years at the start of the experiment **State:** State in which the child resides

Is Male?: Whether the child is male or female. (1: Male 0: Female)

Is Rural?: Whether the child resides in a rural area (1: Rural 0:Non_Rural)

Is Meat Consumer?: Whether the child consumes meat as a part of diet(1: Meat Consumer 0: Vegetarian consumer)

HouseHold Income per month: Income of the household in which the child resides

Base Height: Height of the child in centimetres at the start of the experiment.

Month 1: Height of child at the end of month 1 afier the start of the experiment.

Month 2: Height of child at the end of month 2 afier the start of the experiment.

Month 3: Height of child at the end of month 3 afier the start of the experiment.

Month 4: Height of child at the end of month 4 afier the start of the experiment.

Month 5: Height of child at the end of month 5 afier the start of the experiment.

Month 6: Height of child at the end of month 6 afier the start of the experiment.

Month 7: Height of child at the end of month 7 afier the start of the experiment.

Month 8: Height of child at the end of month 8 afier the start of the experiment.

Month 9: Height of child at the end of month 9 afier the start of the experiment.

Month 10: Height of child at the end of month 10 afier the start of the experiment.

Month 11: Height of child at the end of month 11 afier the start of the experiment.

Month 12: Height of child at the end of month 12 afier the start of the experiment.

Note:

- Analyze data using Jupyter notebook.
- Submit assignments with .ipynb file extension.
- Kindly use comments or markdowns to explain the logic behind the codes.
- Visualize the evidence.
- Write a conclusion.