**Topic 3: Final Min-Quiz:**

**Question 1:**

Upon reviewing your dataset, you determine that responses to an item on weekly number of alcoholic beverages consumed within the past month were more likely to be missing for participants reporting a lower socioeconomic status. What type of missing data is this?

1. Missing at Random (MAR)
2. Missing Not at Random (MNAR)
3. Missing Completely at Random (MCAR)
4. Missing Both at Random (MBAR)

**(Answer: A)**

**Question 2:**

Which of the following is not considered single imputation: (Answer: C)

1. Last Observation Carried Forward
2. Hot-deck
3. Weighting
4. Regression imputation

**(Answer: C)**

**Question 3**

Which of the following best describes a key difference between multiple imputation (MI) and maximum likelihood (ML) imputation?

1. MI creates multiple datasets, while ML estimates parameters directly from incomplete data.
2. ML imputes multiple values per missing observation, while MI fills in missing values only once.
3. MI assumes missing data are missing completely at random (MCAR), while ML does not.
4. ML cannot handle missing data in large datasets, while MI can.

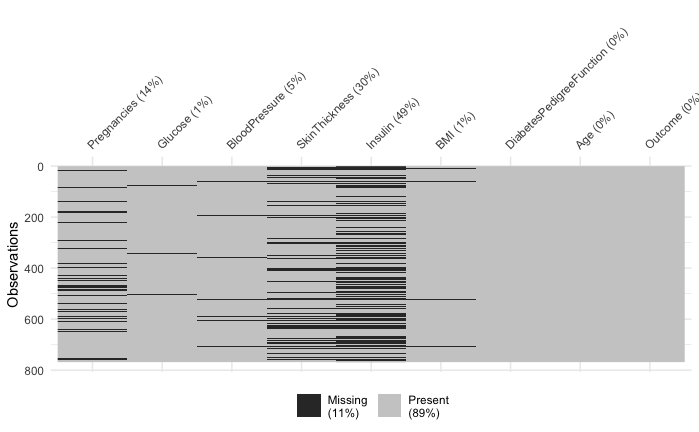
**(Answer: A)**

**Question** **4:**

You run the following R command to visualize missing data:

*vis\_miss(df)*

This is the output. What does this tell you about the dataset?



1. The dataset has no missing values
2. The variables with more black shading have higher proportion of missing data
3. The darker shading represents outliers in the dataset
4. The variable insulin has no missing data

**(Answer B)**

**Question 5:**

When using the mice( ) function in R, what does the m argument specify?

1. The number of iterations the algorithm runs to impute missing values.
2. The number of different imputed datasets created.
3. The number of predictor variables used for imputation.
4. The method used for imputation, such as "pmm" or "norm".

**(Answer B)**