**Mendelian Randomization Exercise: Answers**

**Part 1: Performing the Core MR Analysis**

**Answers for Part 1:**

1. **Based on the mr\_results table, what are the causal effect estimate (b) and p-value (pval) for LDL on CAD for IVW?**
   * **Inverse Variance Weighted (IVW):** Causal effect estimate (b) approx 0.0716, p-value approx 0.1888. **Not significant** at alpha=0.05.

**Part 2: Assessing Heterogeneity**

**Answers for Part 2:**

1. **Examine the forest plot of individual SNP effects. Do all SNPs show effects in the same direction? Do the confidence intervals largely overlap?**

While most SNPs show an effect in the positive direction (i.e., higher LDL leads to higher CAD risk), there is considerable variability in the magnitude of their individual effect estimates. The confidence intervals for the individual SNPs do not all largely overlap; some are quite spread out, and many do not overlap with the overall mean estimate.

1. **What does the spread or consistency of these individual SNP effects tell you about potential heterogeneity in our analysis?**

The wide spread and lack of consistent overlap in the confidence intervals of individual SNP effects strongly suggest the presence of **heterogeneity**.

**MR Egger:** Q = 350.97, Q\_df = 68, Q\_pval = 9.960582e-40

**IVW:** Q = 372.22, Q\_df = 69, Q\_pval = 3.916740e-43

1. **How might significant heterogeneity affect the validity or interpretation of the Inverse Variance Weighted (IVW) method's results?** Significant heterogeneity violates a key assumption of the IVW method: that all instrumental variables estimate the same causal effect. When heterogeneity is present, the IVW estimate can be biased, and its standard errors can be underestimated.

**Part 3: Investigating Pleiotropy with the Funnel Plot**

**Answers for Part 4:**

1. **Describe the shape of the funnel plot. Does it appear symmetrical? What might asymmetry indicate?** The funnel plot does not appear perfectly symmetrical. There is a noticeable spread or lean, particularly at lower precision (towards the bottom of the plot), suggesting that SNPs with less precise estimates tend to deviate more, potentially in one systematic direction. Asymmetry in a funnel plot is a visual indicator of potential **directional pleiotropy**.
2. **Based on the visual evidence from the funnel plot, do you suspect the presence of directional pleiotropy? Why or why not?** Yes, based on the observed asymmetry, particularly the apparent systematic deviation of points at lower precision, we would suspect the presence of directional pleiotropy. If there were no pleiotropy, we would expect a more symmetrical funnel shape, with points scattering evenly around the true causal effect at all levels of precision.

**Part 4: Quantifying Pleiotropy with MR Egger Intercept**

**Answers for Part 4:**

1. **Interpret the egger\_intercept value and its corresponding p-value. What does this statistically tell you about the presence of directional pleiotropy?** The egger\_intercept is approx −0.00296 with a p-value of approx 0.046. Since the p-value is less than 0.05, the MR Egger intercept is statistically significantly different from zero. This provides statistical evidence for the presence of **directional pleiotropy**.
2. **Observe the lines representing the different MR methods on the scatter plot. Which line's intercept is forced through the origin, and which is allowed to deviate? What does this illustrate about the underlying assumptions of these methods regarding pleiotropy?**
   1. The **Inverse Variance Weighted (IVW)** line is forced to pass through the origin (0,0). This visually represents the IVW method's core assumption that there is no directional pleiotropy, meaning any pleiotropic effects of the SNPs on the outcome average out to zero (i.e., the intercept of the regression line is zero).
   2. The **MR Egger** line is allowed to have an intercept that deviates from zero. This demonstrates the MR Egger method's ability to detect and account for directional pleiotropy.
3. **Given the heterogeneity and pleiotropy test results, which MR method’s estimate (IVW or MR Egger do you think is most reliable in this specific scenario, and why?**

Since no pliotropy assumption of IVW is clearly violated. MR Egger would give a more accurate causal estimate.

**Part 6: Leave-One-Out Analysis**

**Answers for Part 6:**

1. **Examine the leave-one-out plot. Are there any individual SNPs whose removal substantially changes the overall causal effect estimate?** In the provided example, while most SNPs individually do not dramatically alter the overall estimate when removed, there might be one or two (e.g., the SNP at the very bottom left of the original forest plot with a strong negative effect, rs1333042) whose removal causes a noticeable shift in the overall estimate. This indicates they are influential outliers.
2. **What could be the implications of an influential SNP in an MR analysis, and what steps might you consider if you identify one?** An influential SNP indicates that a single genetic variant has a disproportionate impact on the overall causal effect estimate. This could imply:
   * A true strong effect, but it might be an outlier.
   * Measurement error in the effect estimates for that specific SNP.
   * Strong pleiotropic effects unique to that SNP.
   * Linkage disequilibrium with an unmeasured causal variant.

If an influential SNP is identified, one might consider:

* + **Investigating the SNP:** Research its biological function, known associations, and potential pleiotropic pathways.
  + **Sensitivity analyses:** Re-running the MR analysis both with and without the influential SNP to assess how much it affects the overall conclusion.
  + **Robust methods:** Reconfirming results with methods more robust to outliers, like the Weighted Median or MR Egger, which you've already done.
  + **Exclusion (with caution):** If there is strong biological or statistical justification (e.g., evidence of strong specific pleiotropy for that SNP), one might justify excluding it.

**Bonus Discussion Question**

1. **Why IVW is not significant for this case, and if MR Egger is underpowered, would you expect the weighted median to perform better in this case?**.
   * **Why not IVW?** The IVW method's core assumption of no directional pleiotropy is violated. In the presence of pleiotropy, the IVW estimate tends towards null.
   * **Why not MR Egger alone?** While MR Egger accounts for directional pleiotropy, it tends to be less precise (having wider confidence intervals) and is highly sensitive to the "InSIDE" (Instrument Strength Independent of Direct Effect) assumption. Its larger estimate might also be less stable with a limited number of instruments.
   * **Why Weighted Median?** The Weighted Median method is more robust to both heterogeneity and pleiotropy (it can provide a consistent estimate even if up to 50% of the weight comes from invalid instruments). Its strong statistical significance (p=1.6^10−5) despite the identified violations suggests a more trustworthy causal effect estimate in this complex scenario, as it has higher power than MR Egger.