* What dataset are we using to answer the question? What are the limitations of the dataset we are using to answer the question?
  + The ANES 2020 Time Series study includes re-interviews with 2016 ANES respondents, a freshly drawn cross-sectional sample, and post-election surveys with respondents from the General Social Survey
  + Respondents assigned to interview by one of three mode groups – by web, video, or telephone
  + 8,280 pre-election interviews and 7,449 post-election interviews
    - *Note: Should filter to individuals for whom both pre-election and post-election results are available (V200004 gives a flag for pre and post-election interviews (both) complete)*
  + 2017 ANES group consisting of respondents who had participated in 2016 Time Series Study
  + Fresh 2020 cross-section
    - Target population: 231 million non-institutional U.S. citizens aged 18 or older living in 50 US states of D.C.
    - Sampling frame: Lists of residential addresses where mail is delivered (from USPS), with all residential addresses across 50 states and Washington DC having equal probability of selection
    - Eligibility: Respondent had to reside at the sampled address and be a U.S. citizen aged 18 or older
  + Limitations:
    - There is potential for the end list of survey respondents to not be nationally representative due to selection bias, but much of this has been accounted for by survey weights
      * There may be some demographic groups that are more likely than others to be ineligible for the survey (e.g., not residing at sampled address)
      * Response rates may differ regionally and by type of dwelling
      * Within each household, there maybe selection bias towards particular individual characteristics who are more likely to respond (e.g., busy, working professionals may have less time to participate in the survey)
    - The data includes both a fresh cross-section and an ANES 2016 Time Series re-interview. There may be differences in the characteristics of both groups – e.g., those who agree in 2016 Time Series re-interview may have different political attitudes and behaviors vs. those in the fresh cross-section
      * To mitigate, it is important to know whether the composition of democrat and republican groups is comparable in terms of these two groups (or stratify results by cross-section vs. 2016 Time Series re-interview cohorts)
      * Notably, the weights applied to the two groups differ, so the 2016 Time Series sample may have additional selection biases that make the sample less nationally representative
* How to operationalize the concept of voter difficulty (how do we measure voter difficulty)?
  + Who or what is a voter?
    - To be considered a voter, use the following variables from the PRE questionnaire:
      * Intend to register to vote / already registered (V201025)
      * Already voted in general election – early vote? (PRE: V201022 / V201023)
      * Cast an actual vote in November 2020 in general election (V202066 – response 4 (I am sure I voted)
  + Who is a ‘Republican’ and who is a ‘Democrat’?
    - [Given] Treat individuals who lean in one direction or another as members of that party e.g., ‘leans Democratic’ classified as Democrat, ‘leans Republican’ classified as Republican
    - Is it reasonable to use the vote that someone cast to identify their party preference in this case? What if someone had so difficult a time voting that they did not cast a ballot?
      * No. Though individuals will typically vote according to their party preferences, there maybe cases where individuals do not have a preference for either party but vote based on their preference of the presidential candidate. Additionally, there maybe groups of individuals who have a preference for Democrat or Republican who did not vote.
    - Recommended operationalisation:
      * PRE-POST: SUMMARY: PARTY OF REGISTRATION (V202065x) – should hold precedence over PRE: SUMMARY: PARTY ID
        + Democrat: 1
        + Republican: 2
      * IF answered ‘None,’ ‘Independent’ or ‘Other for V202065X, use PRE: SUMMARY: PARTY ID (V201231x) -
        + Democrat: 1, 2, 3
        + Republican: 5, 6, 7
      * Include:
        + Democrat / Republican as per V202065x or V201231x
      * Exclude:
        + Did not answer Democrat or Republican in V202065x AND Did not answer Democrat or Republican in 201231x
    - Alternatives would have been to use house of representative, senate, gubernatorial, and/or presidential voting preferences / behaviors (V202105 – V202108), but there are shortcomings to this:
      * One could have conflicting senate, gubernatorial and/or presidential voting preference by party
      * If one were to actually use what vote was actually cast there is an issue of individuals who are ‘Republican’ or ‘Democrat’ who did not vote (e.g., due to difficulty voting)
  + What is difficulty voting?
    - Alternative 1: V202119: How difficult was it for R to vote
      * 5-point Likert scale
      * Benefit is that is most closely approximates the question being answered
      * If we are to assume that this ordinal data behaves like interval data, then a T-test could be used. Would need to check that the distribution is normal without major skew
    - Alternative 2: Sum the counts of ‘1 – mentioned’ for V202120a to V202120j (‘Did R encounter any problems voting’)
      * Result would be metric (interval); T-test could be used
      * Less subjective
      * However, the count of issues voting does not necessarily mean more difficulty voting. For instance, a person could have much difficulty voting but only due to one issue, whereas another could have less difficulty voting but encountered multiple issues in doing so
    - Alternative 3 and 4: V202121 and V202122 (Wait time and how long it takes to get to polling place)
      * Least desirable option as difficulty voting encompasses a much broader set of challenges than just these two issues
    - Could also use composite score, but there are too many issues around weighting
  + Guiding questions for each:
    - What would be the best possible method for measuring this concept? Is this method possible? Why or why not?
    - What is the best available method of measuring this concept?
    - Why have you opted to use this measurement instead of other possibilities? Map the concept definitions that you have written down onto the variables that you are going to use. Describe, precisely, how the variables were generated, if they come from survey data, provide the text of the question that the respondent is reacting to, not the variable name.
    - What, if any, changes have you made to the dataset from how it was provided? Why did you make those changes, how much data was affected, and what are the consequences for any estimates that you produce?
* What is our null hypothesis?
  + [Assuming we do a Mann-Whiteny Test] The probability that difficulty in voting (ranked in a 5-point Likert scale) is higher for Democrats than Republicans is the same as the probability that difficulty in voting (ranked in a 5-point Likert scale) is higher for Republicans than Democrats.
* Which test are we using?
  + Wilcoxon Rank Sum test (Hypothesis of comparison)
    - Assumptions:
      * Ordinal data:
        + Likert scales are ordinal
      * IID data:
        + Both republicans and democrats are drawn from the same distribution – *need to check that the composition of republicans and democrats are equally distributed between the four groups (Cross-Section 2020 (Web-only, Mixed Web, Mixed Video + 2016 ANES). Otherwise, analyse separately*
        + We assume that respondents in the survey are mutually independent. The most likely source of non-independence) would be if multiple members of each household were allowed to conduct the survey (which is not the case since only one potential respondent is drawn from each household – pg. 4).
  + T-test not appropriate as Likert scales are ordinal data and do not meet the assumptions of a t-test
  + Alternatively, chi-squared test could be used but we have used Wilcoxon Rank Sum to focus on content of W203