**Critique of Research Design**

**Overview**

* The overview clarified the main goal of this project objectively, including the situation information. However, an abstract of research approaches, procedures, applied research methods was absent.
* Contrary to the statements in both the Research Question and Treatment sections, which define the goal of this project as to increase turnout in the election and to confirm the increase in the vote propensity caused by the letter service, the goal is defined as to assess if there was an effect of the campaign in this section. Without essential consistency across different sections in this project, the confirmation bias introduced by statements in the later sections could not be limited.
* The concept of participation in the election is raised without specific definition. This caused the conceptual shift from the vote turnout to the vague volunteer engagement or combination of both concepts in the following parts.
* The importance of this research work, its possible contribution to future study, and the generated value or benefit for similar organizations was not clarified. It doesn’t help this research project capture the social attention or funding it deserved. Moreover, background information of the letter-writing campaign was absent without considering diversity of the audience. Introducing the name of “TurnOutForWhat!” without necessary clarification, more confusion was raised from the audience.

**Research Question:**

* First, the research question that is outlined here must be the focus of the remainder of the proposal. That mostly holds true with the central question being around increasing voter participation, but in the Treatment section, there is the second goal outlined of “to increase political engagement among the volunteers in your organization.” Is it actually both and the 2nd is meant to be a sub-question here? If so, that secondary goal must be called out and subsequently discussed throughout the proposal. If the research question really is just around increasing voter participation among the recipients (this is the course of action recommended by the review team), then the focus of this section is good.
* The research question is a bit unclear as it focuses on the action of “letters sent via USPS.” Are you studying an effect of the act of sending a letter making the sender more likely to vote, or on the act of receiving a letter making the recipient more likely to vote (or is it both as mentioned above)? Either way, make this more clear.
* It does not seem that USPS is central to the study design, and can likely be omitted. Sending letters through a different courier would allow for functionally the same study.
* Is there some reason that the 2020 General Election in CA districts 14 and 15 was chosen? Is the thought that seeing an effect in this district in this election would translate to a possible wider implication of the effect of receiving a letter? Defining the question here correctly is important.
* The variable response here needs to be rigidly defined. If you are just looking at the 2020 General Election in CA 14/15, it is perhaps better to say “increase voter participation” rather than “one’s propensity to vote.” Propensity to vote lends itself to be defined as some subjective likelihood of voting. Are you assessing whether voters themselves say whether they are likely to vote? By the rest of the proposal this seems doubtful, as if you’re more concerned with who actually does vote.
* Lastly, in later parts of this proposal it seems as though there is a desire to see some positive effect of letter writing on voter participation (by mentioning tailoring the letters to the recipients in the treatment section), yet this question is written as if this will be an objective study of participation. I’m cautious of confirmation bias. If this is going to be as objective as stated here, extra precautions may need to be taken to prevent bias (ensuring random sampling, thorough and unbiased review, etc.)

**Data**

* The memo is not very detailed on what the data source looks like or where it will come from. What exactly will make up the registration data? Names and addresses, or more detailed data like political party association and previous voting history?
* Why are only the current and immediately previous elections considered in the definition of likely voters? Wouldn’t someone with a pattern of voting in many previous elections be more likely to vote in the next one as long as they are still alive and live in the relevant district?
* There are issues with using voter registration data to define likely voters. For example, how can you know that someone is likely to vote for the first time in 2020? Just because someone has registered to vote, does not necessarily imply that they will (especially if the district employs automatic voter registration). A second concern is that voters who voted in the 2016 election in the examined districts may have moved. The ability to extract likely voters just from registration data seems limited.
* Extracting a representative group of likely voters from registration data is subject to sampling bias. If not carefully selected, the sample group could contain all new voters on one extreme, and all existing voters on the other extreme.
* Will there be any need to transform raw data, or clean it in anyway? How will raw data from voter registration be transformed into some sort of measure of likeliness to vote?

**Experimental Design**

* Somewhat vague definition of “starting with two random letters.” Does that mean that two random letters are getting selected, and then all last names that start with the first random letter and then the second random letter consecutively are part of the control group? Or two letters are selected and the set of last names that start with either of those letters are part of the control group. Seems to lean towards the second definition but could use more clarification.
* Choosing only two random letters for last names will result in a treatment group that is much bigger than the control group. Shouldn’t the groups be about the same size?
* Randomly picking a letter for last names can bias towards certain ethnic and socioeconomic subgroups of a district’s population, potentially resulting in misestimation of a districts tendencies to vote.
* Should focus more on trying to send letters to subsets of the population who are on the fence about voting, and then split that group into a control and treatment group. Sending letters to a randomly selected population could result in primarily sending letters to a group of voters who would already more than likely vote.
* How can it be confirmed that the person who was sent a letter actually received it and opened it?
* What exactly will the letter say? Could use more clarification about the actual contents of the letter and why those contents were chosen.
* Will the letter be sent in multiple languages? If a large portion of the treatment group does not speak English particularly well and the letter is only provided in English, then it could be concluded that the letter had or didn’t have an effect when it was simply ignored.

**Statistical Methods:**

* There is not enough information for an average reader to understand about the proposed statistical methods.
* More details need to be provided around what kind of conventional statistical methods to be used. A better way would be to provide details about what statistical test is used like T-test or any other test.
* Will accounting for demographic factors be done in the experimental design??
* How will they account for demographic factors? Wouldn’t it be easier to account for those same factors up front and ensure that each group (treatment and non-treatment) includes a representative sampling of each?
* Qualitative analysis vs quantitative analysis: Qualitative analysis about demographic factors is based on literature review, however, it requires collecting information about age, previous voting history, and race of study group. Research method of this part is missing
* Does this lead to any anchor biases to start off? Using conventional PS methods for voter participation rates and accounting only demographic factors that influenced turnout in the past could lead to depending on only a set of conventional PS methods.

**Sample Size**

* To guarantee the representative and randomness of the sample, a sufficiently large sample size should be determined by calculation when it is possible. In this project, the sample size was claimed to be decided by statistical analysis. There is an absence of information of the applied statistics model, calculation method, or reference of previous research sample size in the same field. However, it was more related to the population of selected family names according to the experiment design part, which is of no statistical analysis guarantee.
* According to the Treatment part, the actual sample size was definitely exaggerated without counting the return mail. In this proposal, the sample size is actually decided by the number of the letters sent. Without a proper definition of sample, e.g. letter receiver who does read it, this increases the risk of small sample size. Need a way to find out a true estimate of sample size that will provide an accurate estimate of the participation of voters
* Due to the confirmation bias and anchoring bias introduced by the research question, the bias that “enough to account for the increase of voting propensity” is more likely to exist in the estimation of proper sample size.
* Overemphasizing the reflection of organization budget, which implies that the budget overweights the importance of sufficient sample size. Actually, the budget plan should be made the best use to accomplish the designed sample size, instead of sacrificing sample size to satisfy budget saving requirements.
* Without considering population information of District 14, 15, it is challenging to decide an applicable experimental group size. Moreover, validity of the experiment design is impossible to be confirmed without the comparison of control group and experimental group size.

**Treatment:**

* As mentioned in the research question section “Increasing political engagement of volunteers” isn’t mentioned anywhere else. So if this is going to be a goal it needs to be called out at the beginning of the proposal and it needs to be defined how this will be measured (surveys, interviews, other qualitative methods, etc.) and what the metrics for impact are.
* The letter timing should be defined. Consulting with USPS seems like a reasonable step to ensure the letters arrive in time, but this should be done now, and a timeline aligned on. Saying “the letters will be sent out 2 weeks in advance, to ensure they arrive on time but also so that they are received close to the time of the election” would suffice.
* A return address should be included to see who will not have received the letter and to understand who to remove from the “letter recipient” population. Otherwise, if a certain effect is seen in a particular demographic, a conclusion might erroneously be “A noticeable effect was seen among the 65+ demographic, suggesting older people are more informed by letters/mail” when really the older population may just not move as much and were more likely to get the letter.
* Each person receiving a very different letter is worrying, as it introduces another variable. Then the skill of the writer (how persuasive they are and how they can connect to the recipient) plays a large role. But this isn’t about “who is the best letter writer” but “does receiving a letter have an effect.” So perhaps receiving a more uniform letter would be advisable.

**Deliverables Post-Election:**

* Timeline : The timelines are not mentioned in detail. It would have been more helpful if the timeline for each research stage was highlighted in detail. Even if it is not a perfect one, an approximation of the magnitude would be useful.
* Deliverables : Deliverables are vaguely mentioned. It seems like the deliverables will be the data analysis report and a result set of the statistical significance test saying there is some effect.