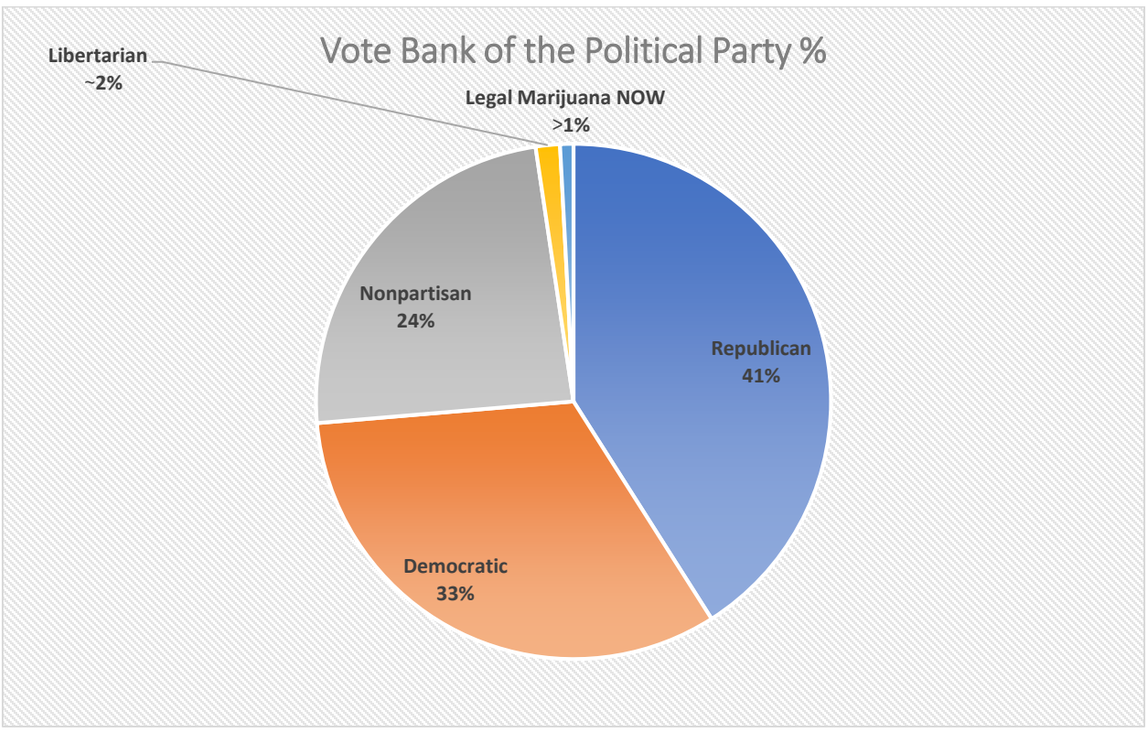
Weighted Survey Design analysis

Weighted survey design analysis is a statistical technique used to adjust survey data to better represent a target population, correcting for sampling biases, nonresponse, or unequal selection probabilities. In the context of dataset, there are four excel sheets ("Full Target Population," "Voter Guide Population," "Survey Population," and "Survey Response Raw Data"), this involves using the survey or srvyr packages in Rstudio to perform:

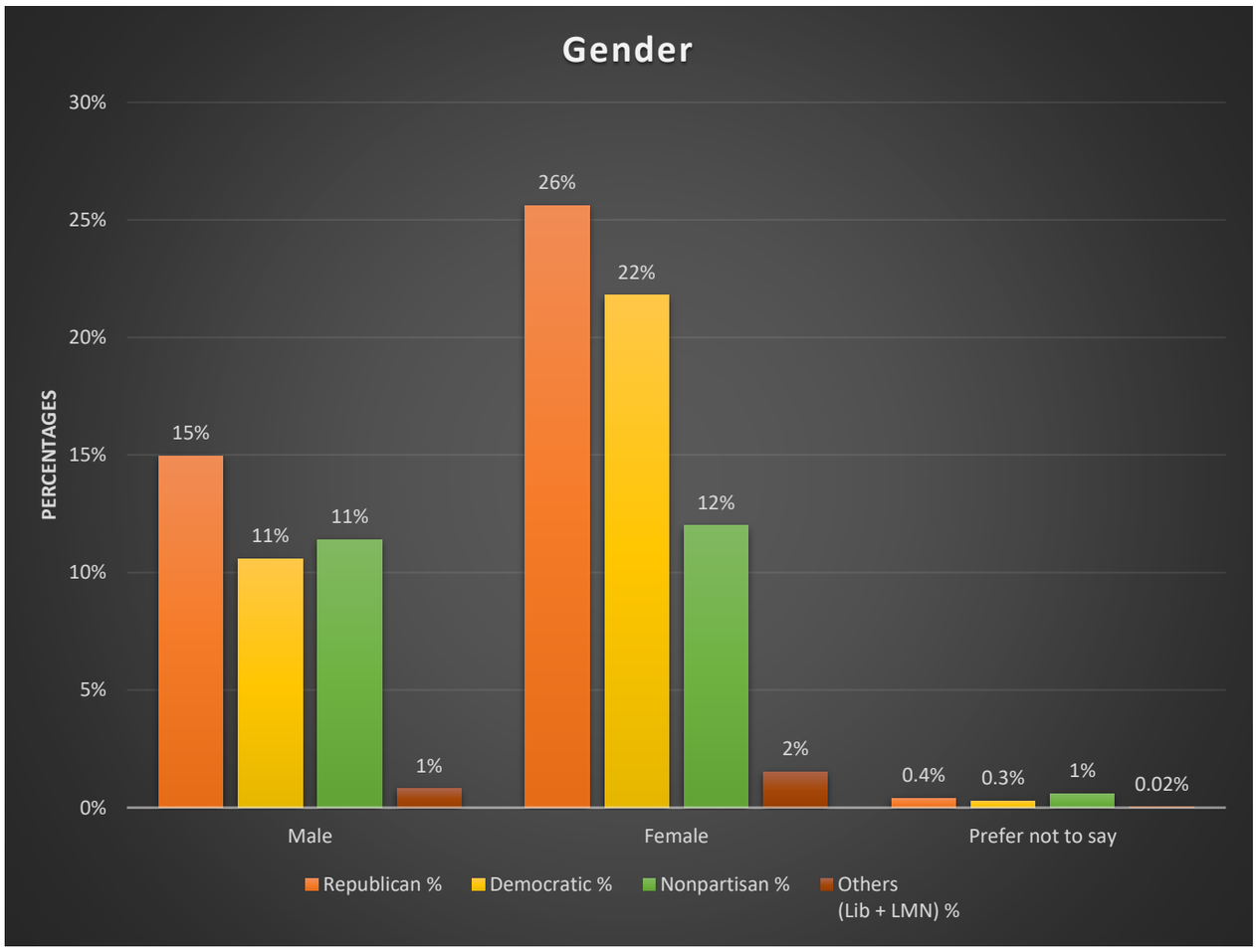
- Weight the 248 survey responses to align with the demographics of the 99830 target population (based on variables like Voter Political Party).
- Performed nonresponse analysis to assess bias by comparing respondents to nonrespondents.
- Generate weighted statistics (toplines, crosstabs) and visualizations.

Below, the results in the following tables and crosstabls showed the weighted estimated results based on the Survey Response Raw Data.

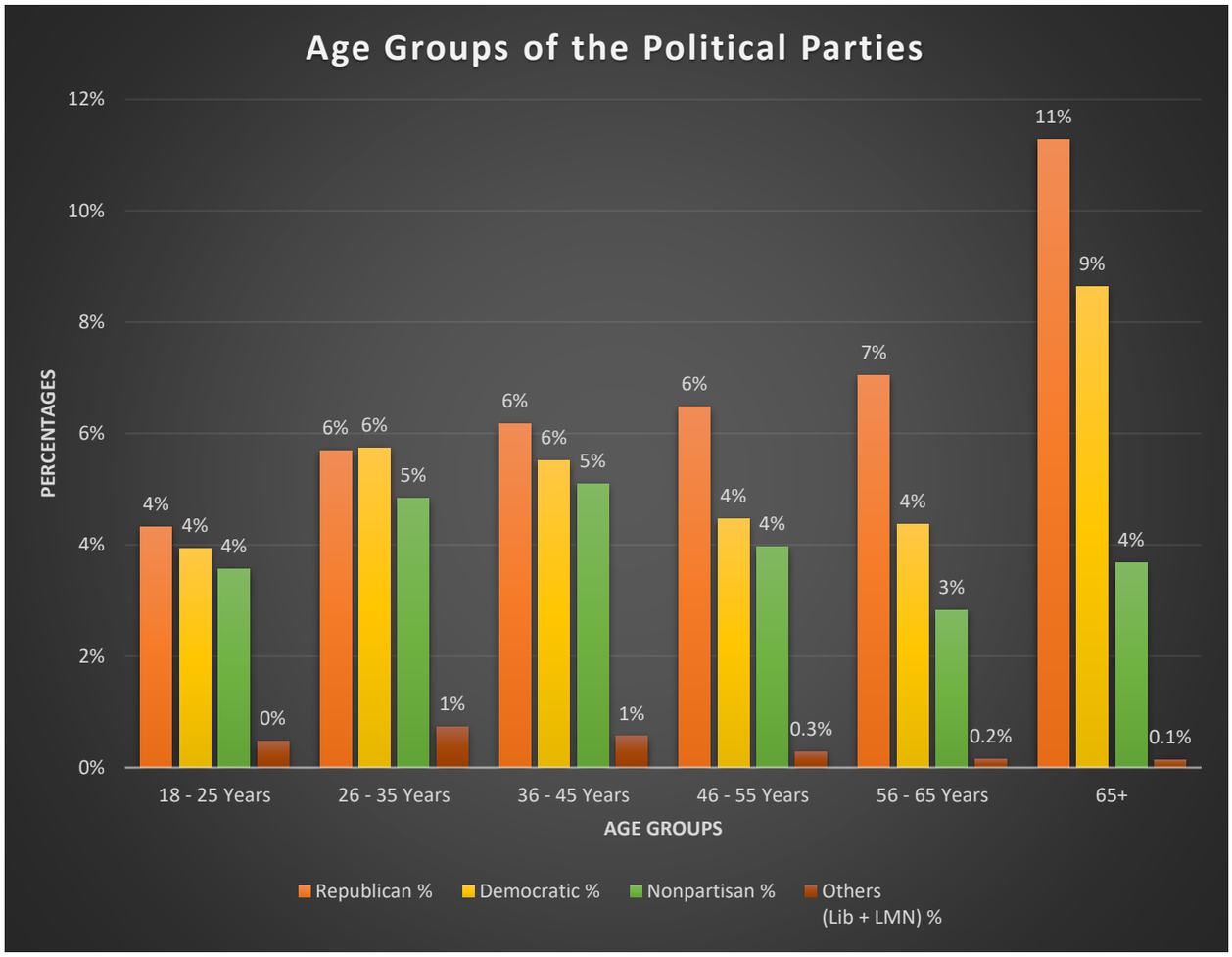
Voter Political Party		
Political Party Name	**Number of Voters**	**%**
Republican	40919	41%
Democratic	32614	33%
Nonpartisan	23944	24%
Libertarian	1518	~2%
Legal Marijuana NOW	835	>1%
Total	**99830**	**100%**



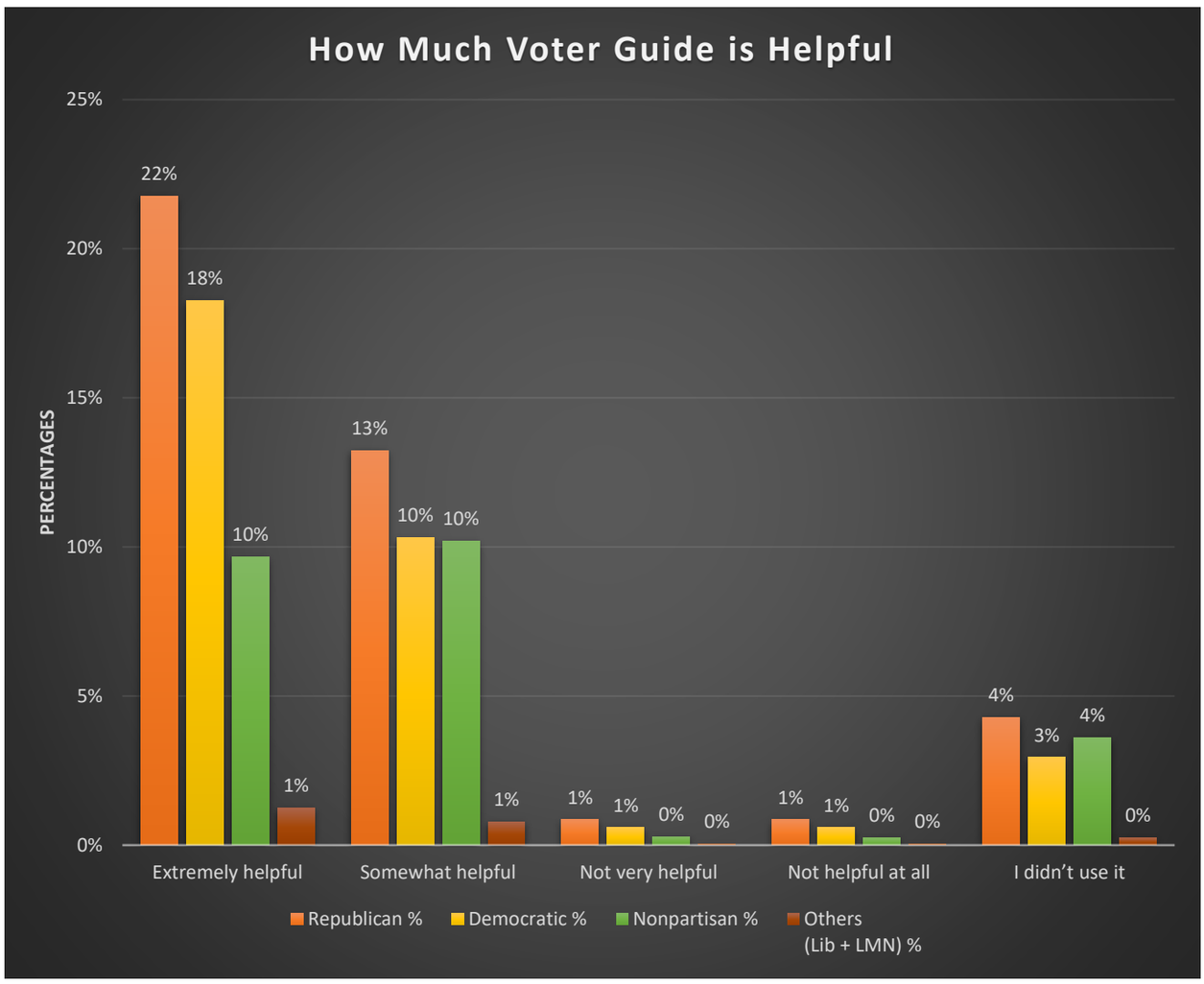
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Sex*	**Male**	14935	15%	10567	11%	11373	11%	800	1%	**37676**	**38%**
	Female	25574	26%	21754	22%	11972	12%	1529	2%	**60830**	**61%**
	Prefer not to say	409	0.4%	294	0.3%	599	1%	24	0.02%	**1325**	**1%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



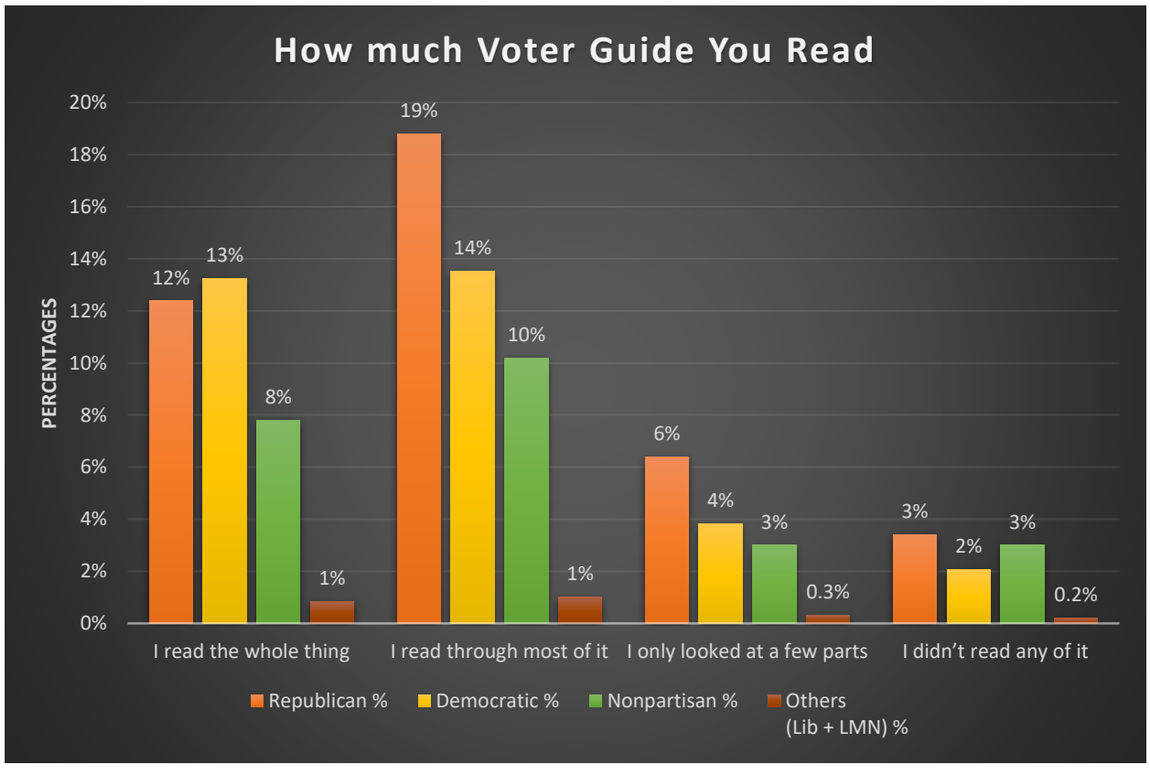
	Groups	Voter Political Party									
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)		Total	
		N	%	N	%	N	%	N	%	N	%
Age in Years	18 - 25	4323	4%	3931	4%	3566	4%	473	0%	**12293**	**12%**
	26 - 35	5688	6%	5731	6%	4833	5%	736	1%	**16988**	**17%**
	36 - 45	6161	6%	5512	6%	5087	5%	564	1%	**17324**	**17%**
	46 - 55	6463	6%	4461	4%	3962	4%	285	0.3%	**15171**	**15%**
	56 - 65	7024	7%	4359	4%	2823	3%	152	0.2%	**14358**	**14%**
	65+	11260	11%	8620	9%	3673	4%	143	0.1%	**23696**	**24%**
Total		**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



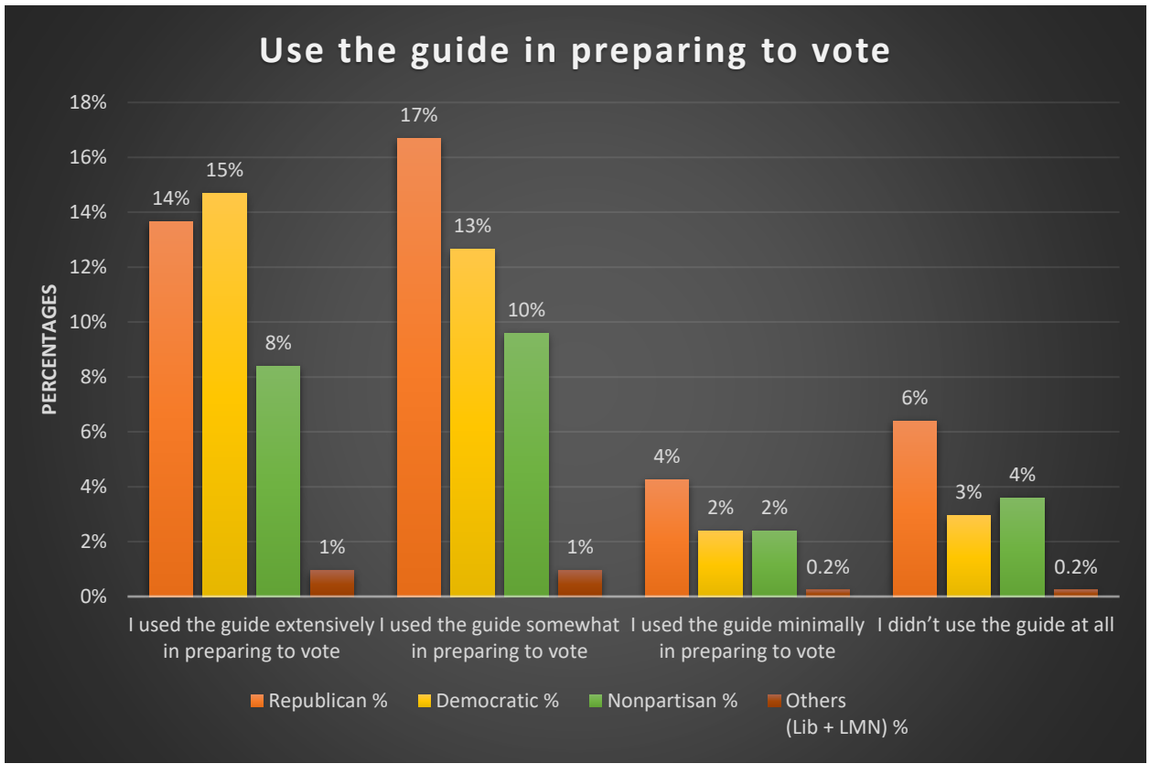
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q1 Voter Guide Helpful	**Extremely helpful**	21738	22%	18217	18%	9649	10%	1243	1%	**50848**	**51%**
	Somewhat helpful	13213	13%	10284	10%	10176	10%	787	1%	**34461**	**35%**
	Not very helpful	852	1%	588	1%	287	0%	38	0%	**1765**	**2%**
	Not helpful at all	852	1%	588	1%	239	0%	38	0%	**1718**	**2%**
	I didn't use it	4262	4%	2938	3%	3592	4%	247	0%	**11039**	**11%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



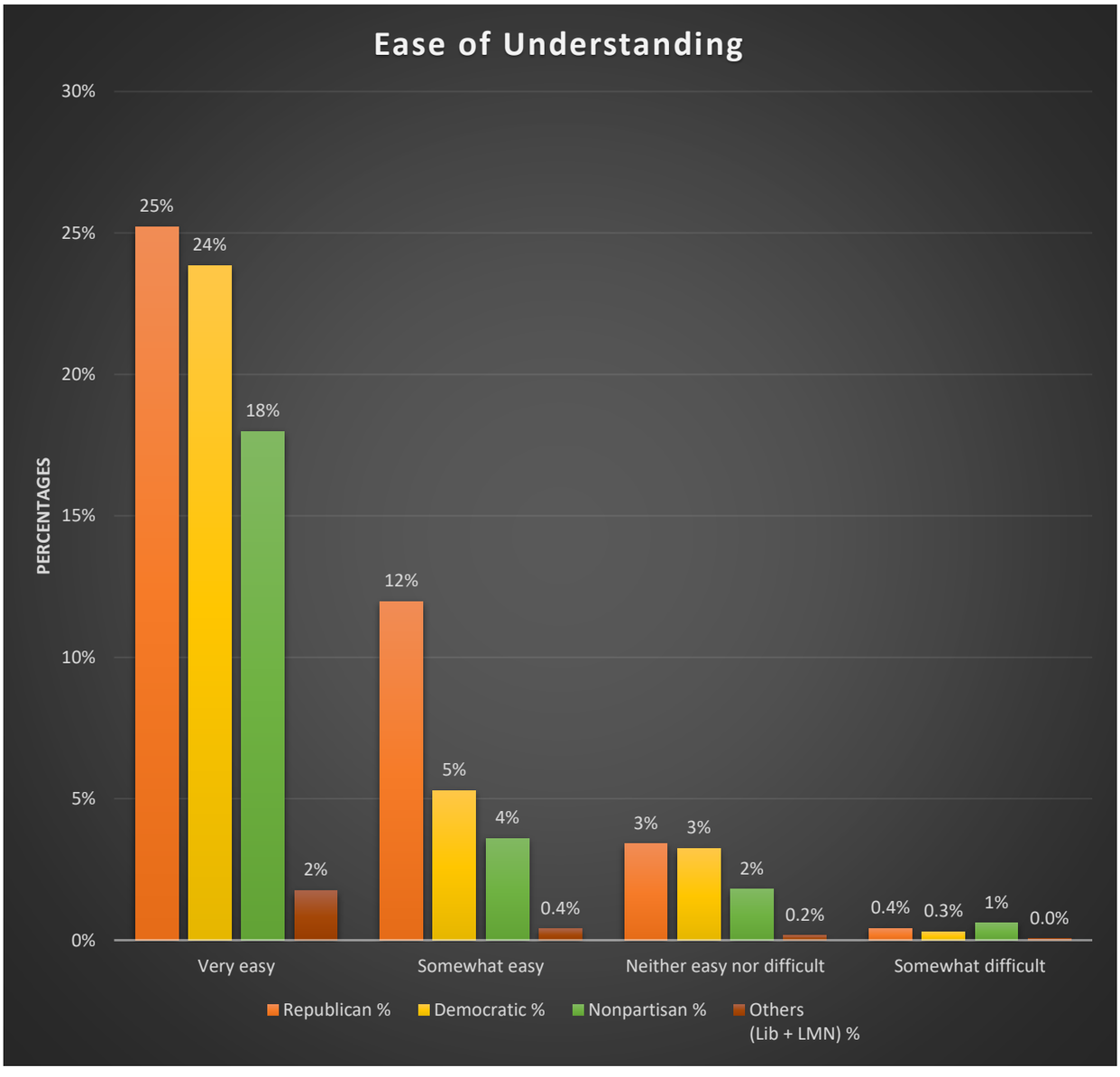
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q2 How much Voter Guide You Read	I read the whole thing	12361	12%	13222	13%	7782	8%	835	1%	**34200**	**34%**
	I read through most of it	18755	19%	13516	14%	10176	10%	1015	1%	**43462**	**44%**
	I only looked at a few parts	6394	6%	3820	4%	2993	3%	313	0.3%	**13519**	**14%**
	I didn't read any of it	3410	3%	2057	2%	2993	3%	190	0.2%	**8649**	**9%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



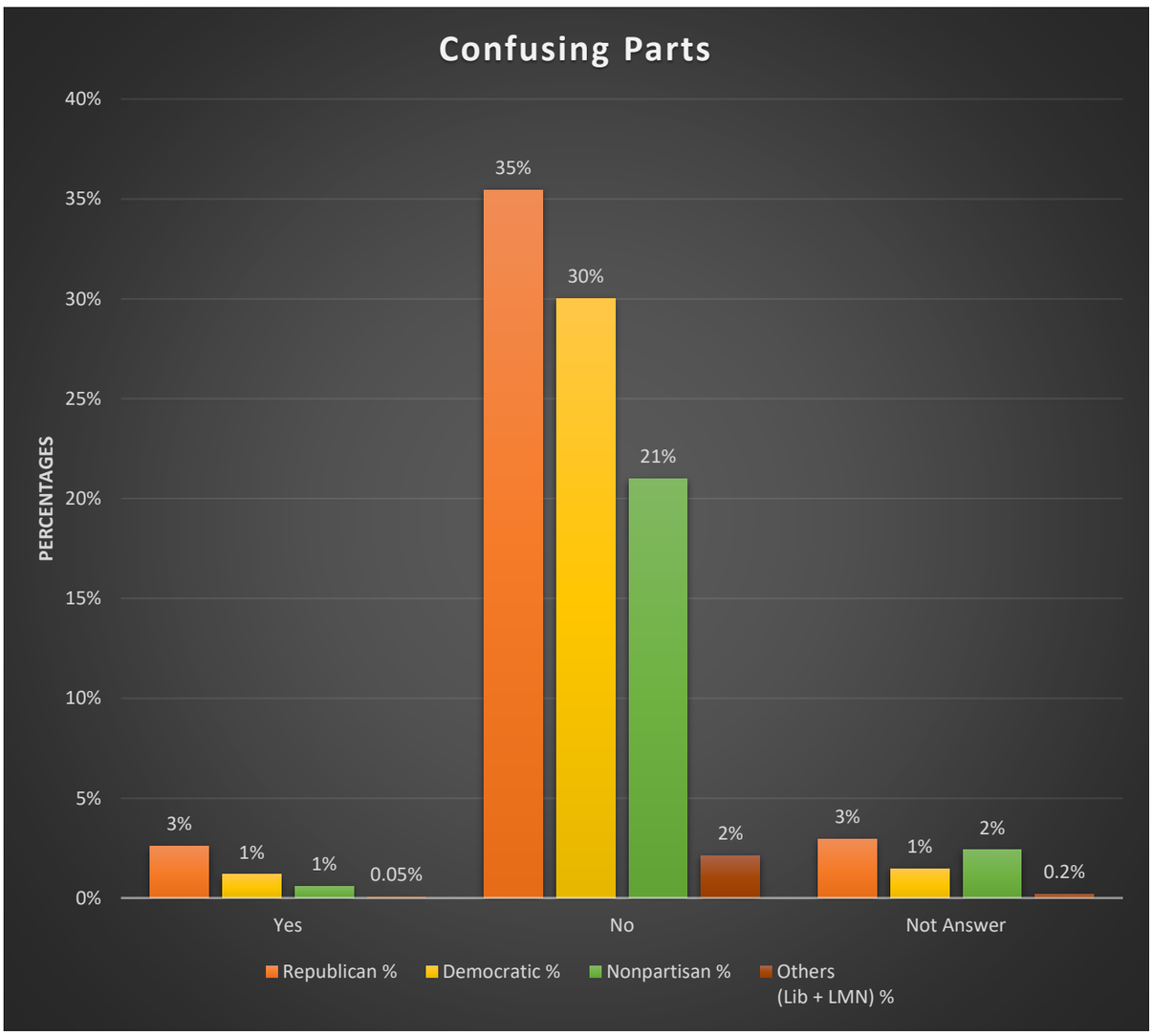
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q3 Usage Preparing	I used the guide extensively in preparing to vote	13626	14%	14676	15%	8380	8%	941	1%	**37624**	**38%**
	I used the guide somewhat in preparing to vote	16654	17%	12622	13%	9578	10%	941	1%	**39795**	**40%**
	I used the guide minimally in preparing to vote	4256	4%	2381	2%	2394	2%	235	0.2 %	**9266**	**9%**
	I didn't use the guide at all	6383	6%	2935	3%	3592	4%	235	0.2 %	**13146**	**13%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



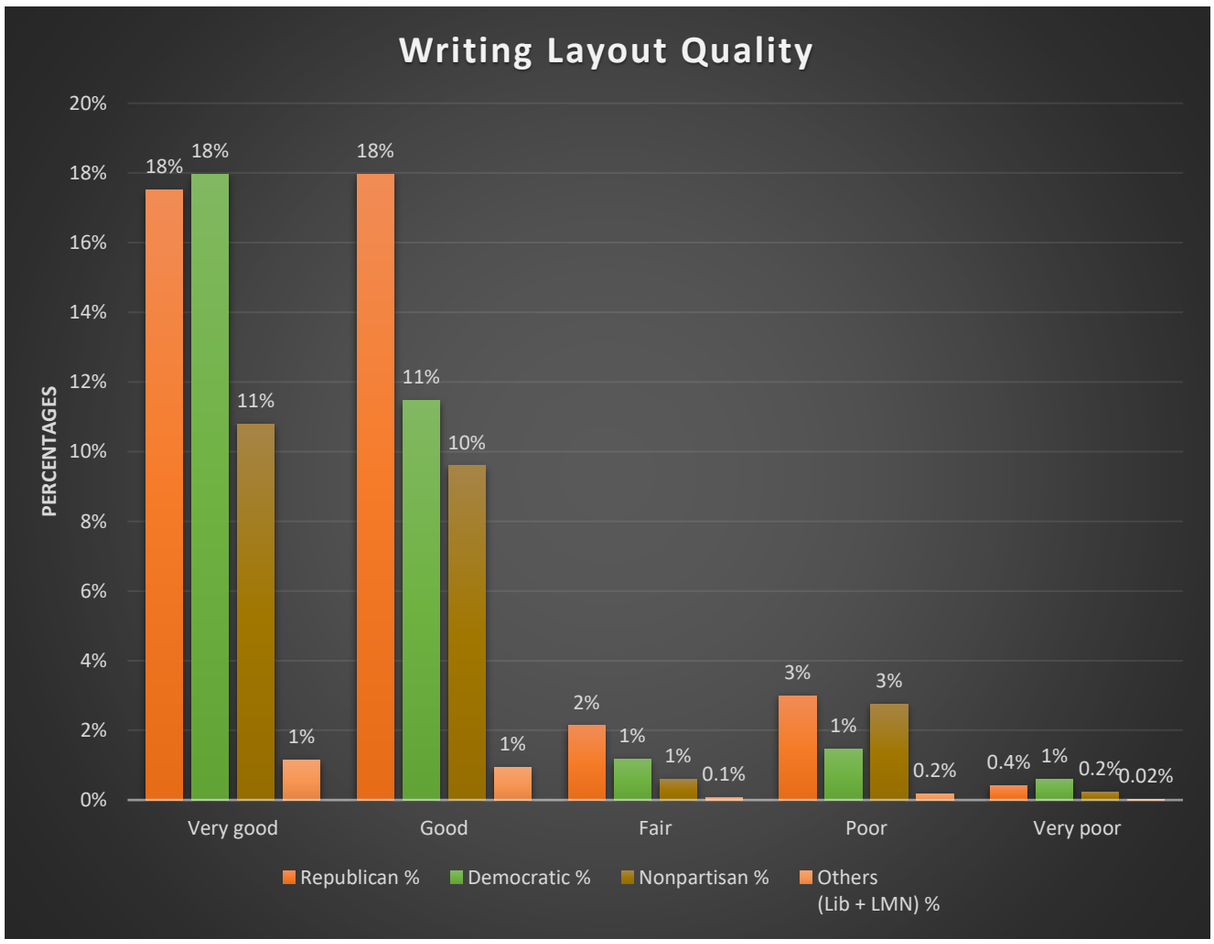
Groups		Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q4 Ease of Understanding	Very easy	25165	25%	23808	24%	17958	18%	1741	2%	**68673**	**69%**
	Somewhat easy	11948	12%	5283	5%	3592	4%	400	0.4%	**21224**	**21%**
	Neither easy nor difficult	3396	3%	3229	3%	1796	2%	165	0.2%	**8586**	**9%**
	Somewhat difficult	409	0.4%	294	0.3%	599	1%	47	0.0%	**1348**	**1%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



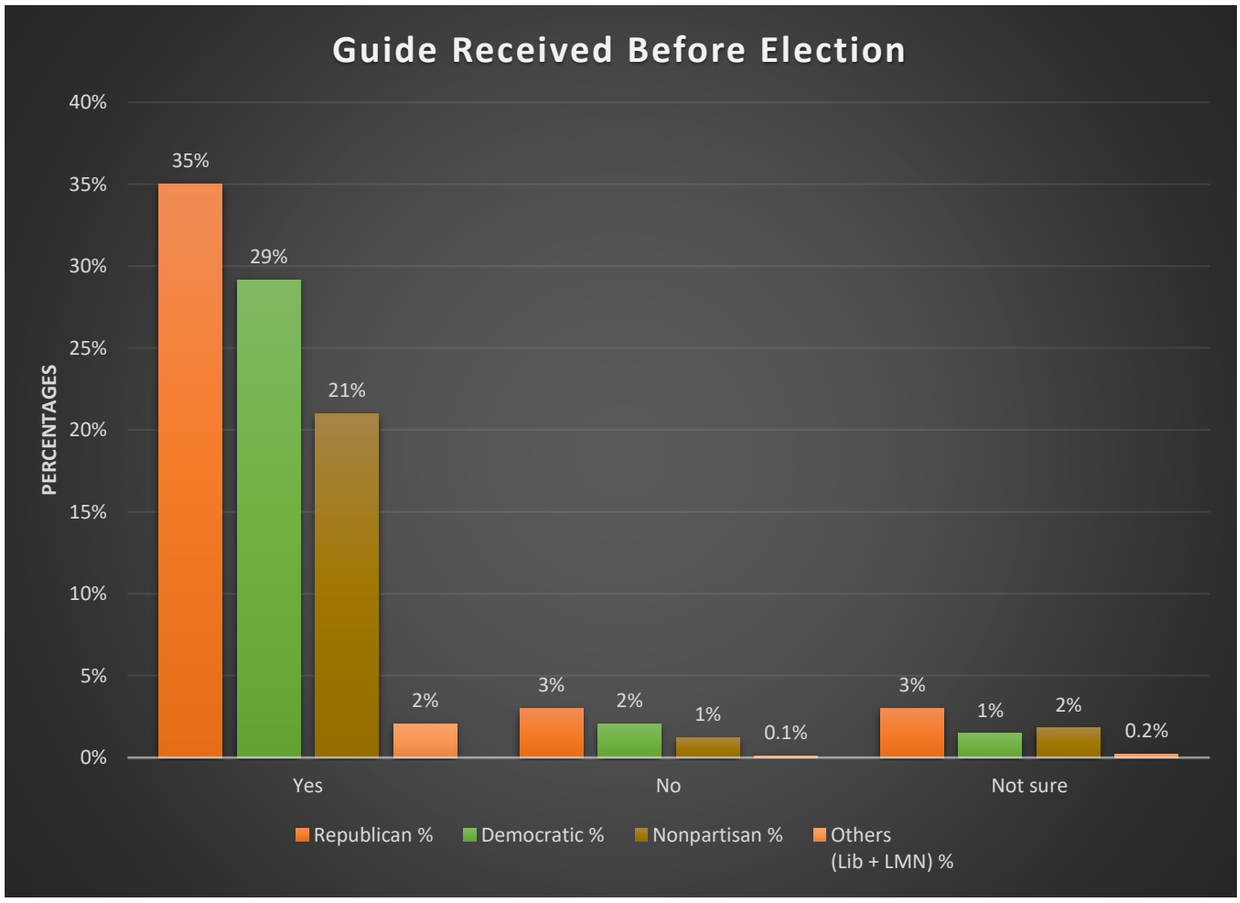
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q5 Confusing Parts	Yes	2578	3%	1174	1%	599	1%	47	0.05%	**4398**	**4%**
	No	35395	35%	29972	30%	20951	21%	2118	2%	**88437**	**89%**
	Not Answer	2946	3%	1468	1%	2394	2%	188	0.2%	**6997**	**7%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



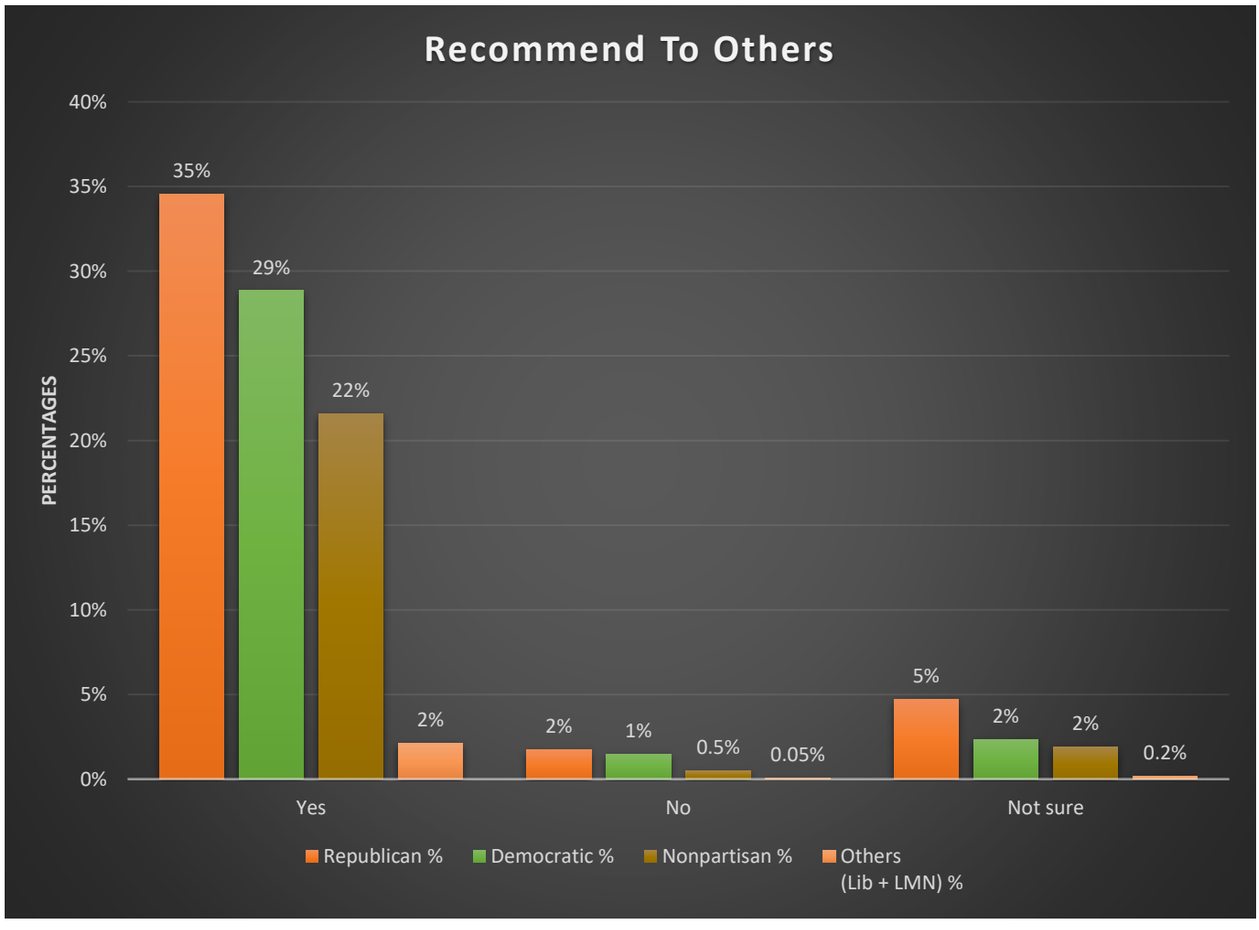
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q6 Writing Layout Quality	**Very good**	17472	18%	17938	18%	10775	11%	1153	1%	**47338**	**47%**
	Good	17923	18%	11448	11%	9578	10%	941	1%	**39889**	**40%**
	Fair	2128	2%	1174	1%	599	1%	71	0.1%	**3971**	**4%**
	Poor	2987	3%	1468	1%	2754	3%	165	0.2%	**7373**	**7%**
	Very poor	409	0.4%	587	1%	239	0.2%	24	0.02%	**1259**	**1%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



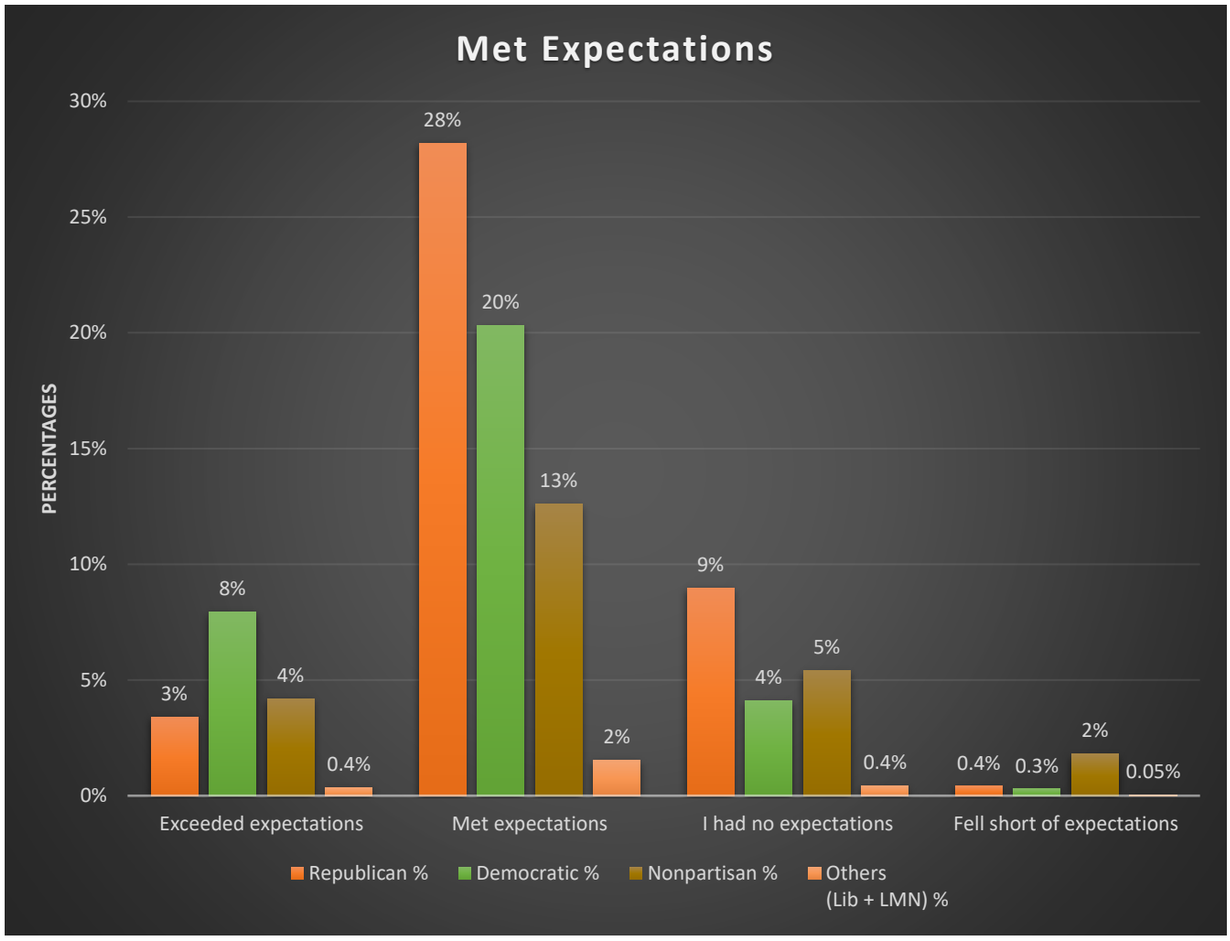
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q7 Receive Before Election	Yes	34945	35%	29092	29%	20951	21%	2071	2%	**87059**	**87%**
	No	2987	3%	2055	2%	1197	1%	94	0.1%	**6333**	**6%**
	Not sure	2987	3%	1468	1%	1796	2%	188	0.2%	**6439**	**6%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



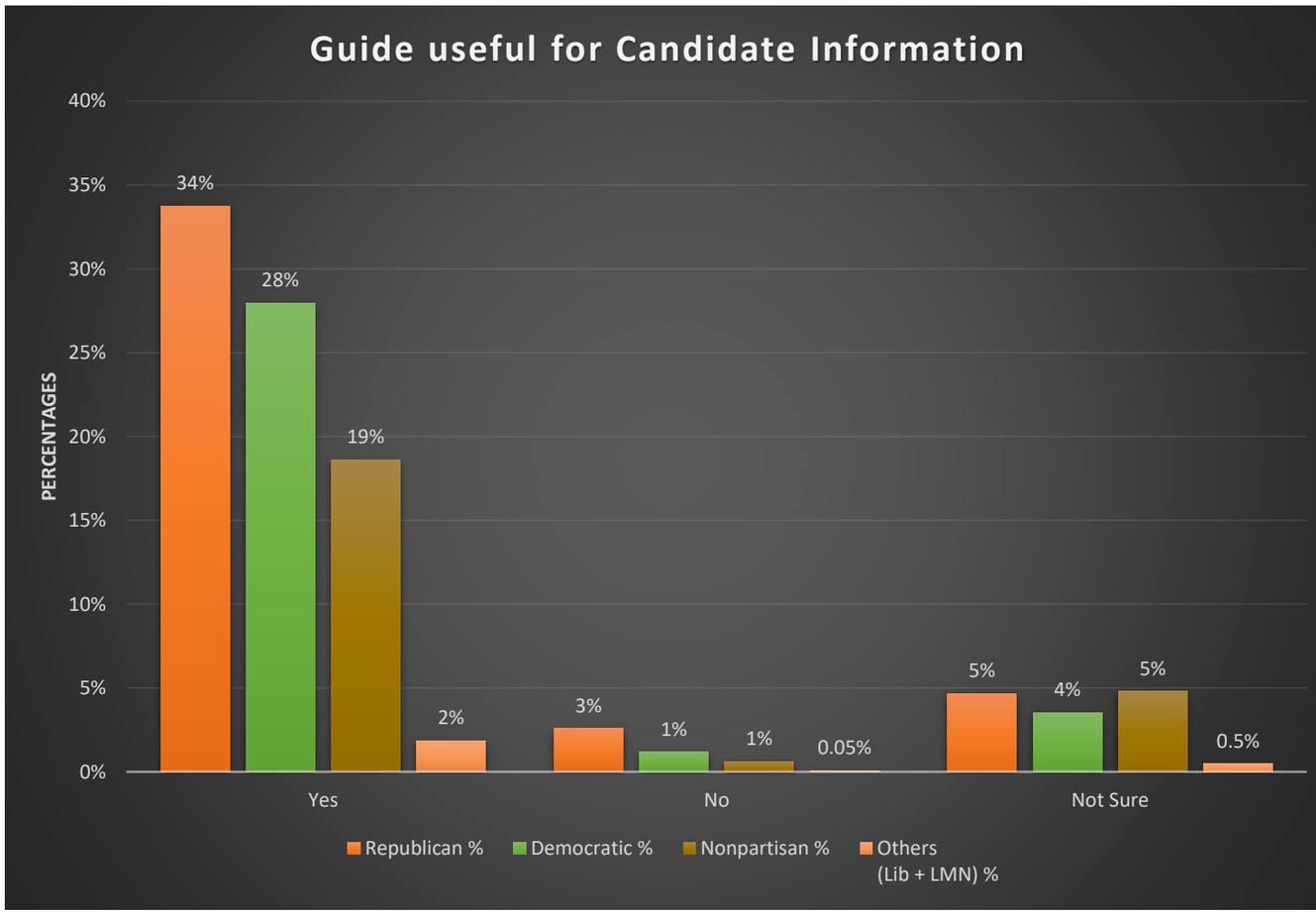
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q8 Recommend To Others	**Yes**	34495	35%	28798	29%	21550	22%	2118	2%	**86961**	**87%**
	No	1719	2%	1468	1%	479	0.5%	47	0.05%	**3712**	**4%**
	Not sure	4706	5%	2348	2%	1916	2%	188	0.2%	**9158**	**9%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



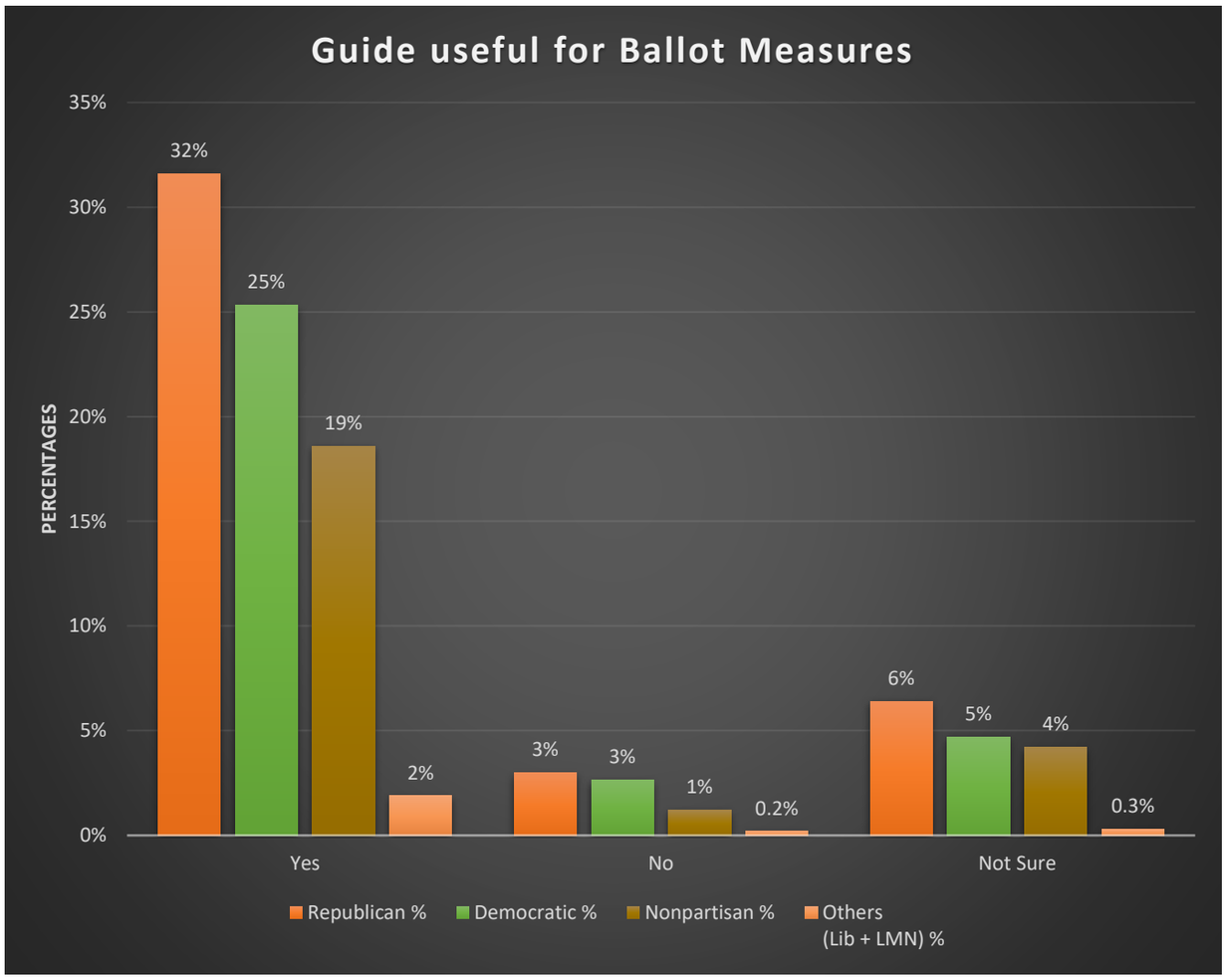
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q9 Met Expectations	**Exceeded expectations**	3396	3%	7925	8%	4190	4%	353	0.05%	**15865**	**16%**
	Met expectations	28152	28%	20286	20%	12571	13%	1529	2%	**62539**	**63%**
	I had no expectations	8961	9%	4109	4%	5387	5%	424	0.4%	**18882**	**19%**
	Fell short of expectations	409	0.4%	294	0.3%	1796	2%	47	0.4%	**2546**	**3%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



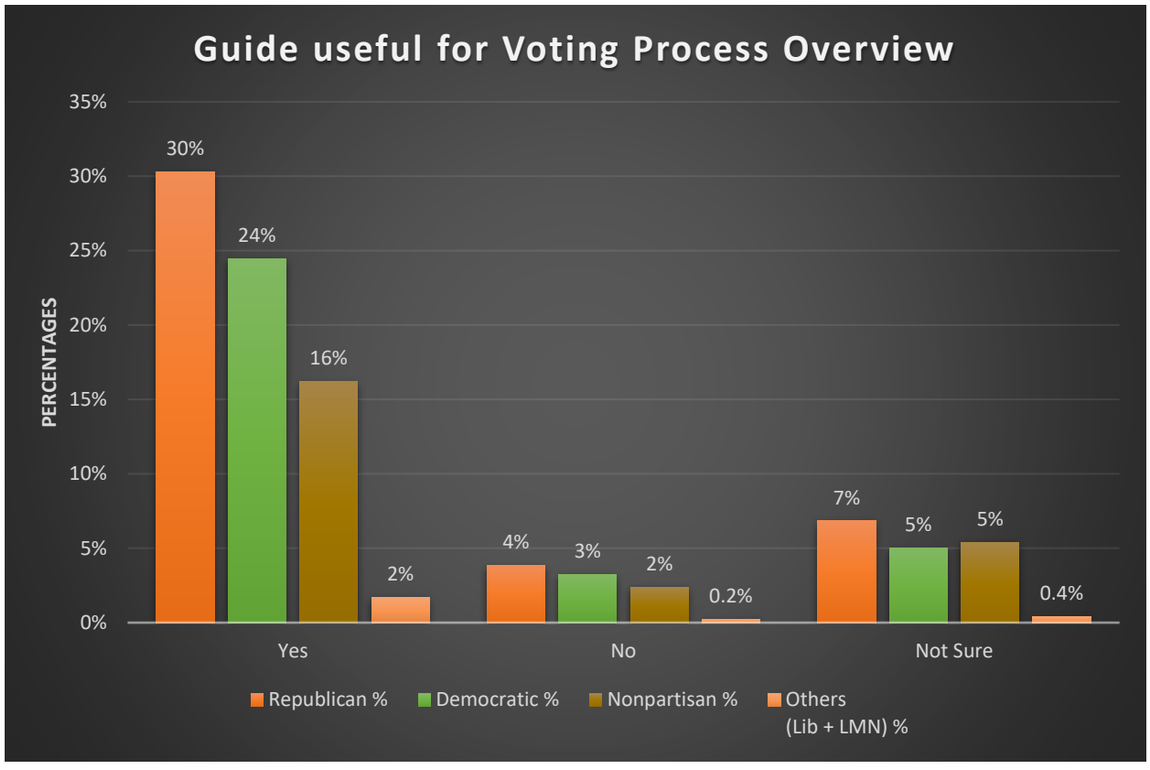
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q10a_Candidate info_Useful	Yes	33676	34%	27918	28%	18557	19%	1835	2%	**81987**	**82%**
	No	2578	3%	1174	1%	599	1%	47	0.05%	**4398**	**4%**
	Not Sure	4665	5%	3522	4%	4789	5%	471	0.5%	**13447**	**13%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



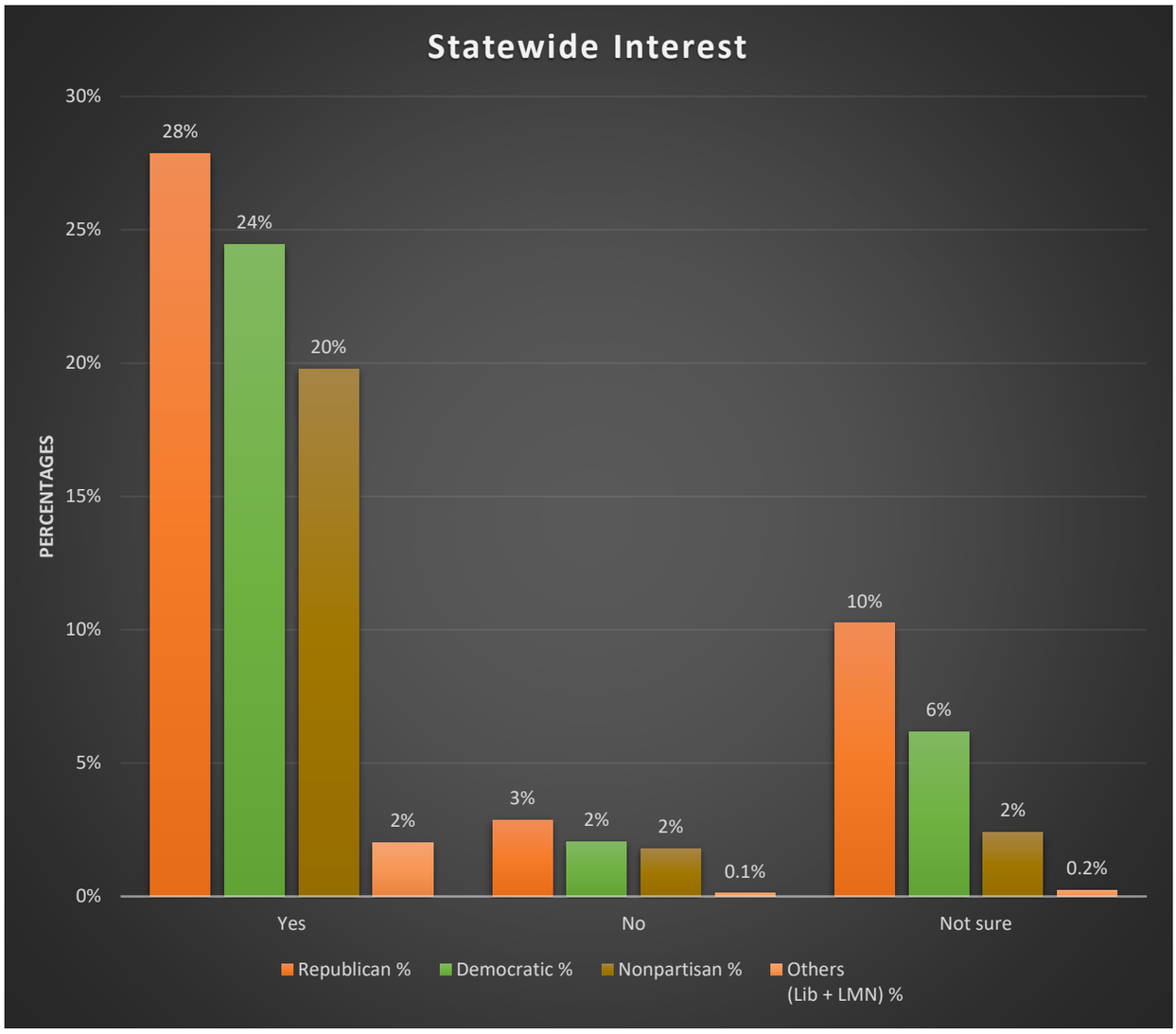
Q10b Ballot Measures Useful	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
	Yes	31549	32%	25276	25%	18557	19%	1882	2%	**77264**	**77%**
	No	2987	3%	2642	3%	1197	1%	188	0.2%	**7014**	**7%**
	Not Sure	6383	6%	4696	5%	4190	4%	282	0.3%	**15552**	**16%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



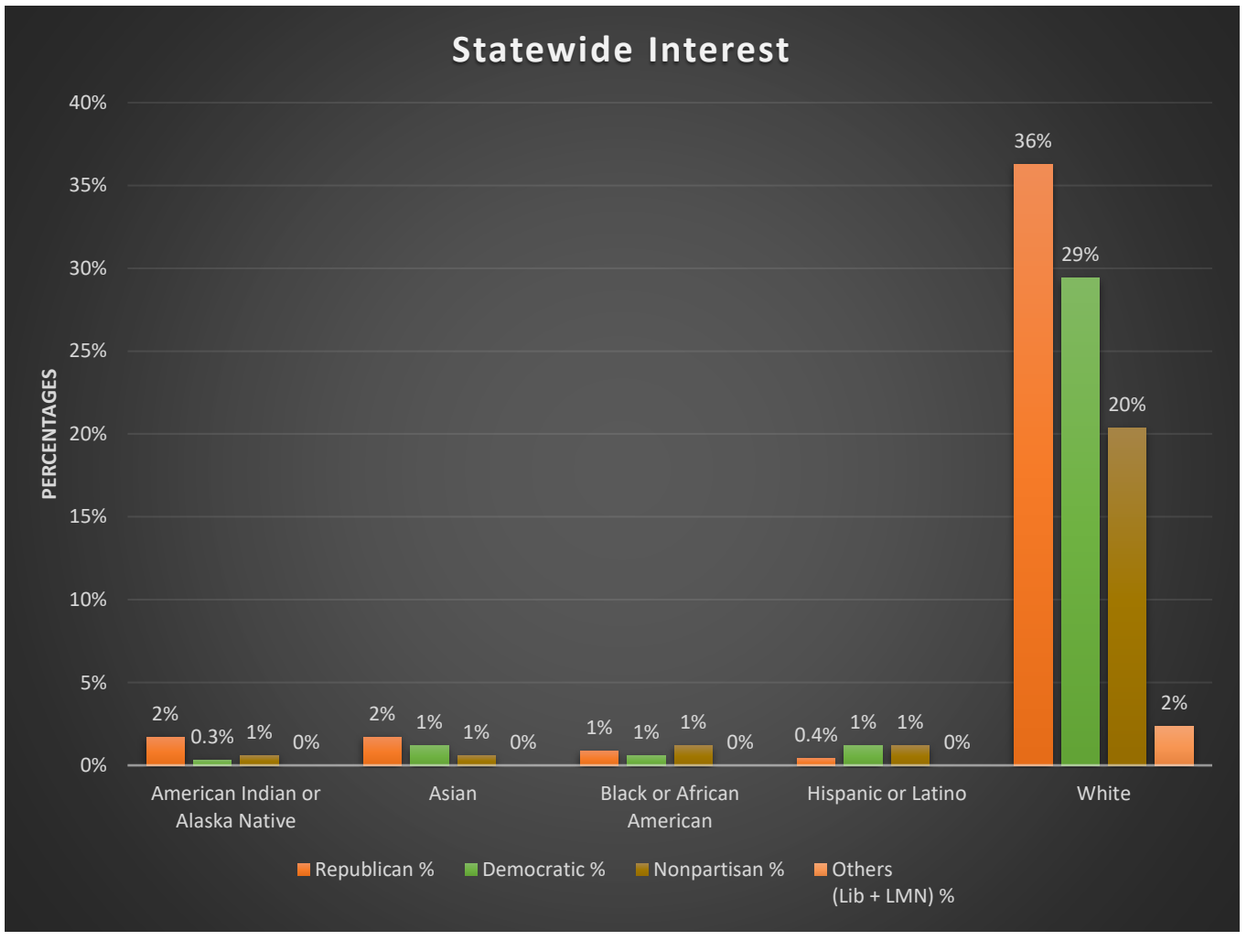
	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q10c Voting Process Overview Useful	**Yes**	30263	30%	24387	24%	16162	16%	1718	2%	**72531**	**73%**
	No	3836	4%	3232	3%	2394	2%	235	0.2%	**9698**	**10%**
	Not Sure	6820	7%	4995	5%	5387	5%	400	0.4%	**17602**	**18%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



Q12 Statewide Interest	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
	Yes	27825	28%	24395	24%	19754	20%	2000	2%	**73975**	**74%**
	No	2864	3%	2055	2%	1796	2%	118	0.1%	**6833**	**7%**
	Not sure	10230	10%	6164	6%	2394	2%	235	0.2%	**19024**	**19%**
	Total	**40919**	**41%**	**32614**	**33%**	**23944**	**24%**	**2353**	**2%**	**99830**	**100%**



	Groups	Voter Political Party								Total	
		Republican		Democratic		Nonpartisan		Others (Lib + LMN)			
		N	%	N	%	N	%	N	%	N	%
Q14 Race Ethnicity	American Indian or Alaska Native	1705	2%	326	0.3%	599	1%	0	0%	**2630**	**3%**
	Asian	1705	2%	1175	1%	599	1%	0	0%	**3479**	**3%**
	Black or African American	852	1%	588	1%	1197	1%	0	0%	**2637**	**3%**
	Hispanic or Latino	426	0.4%	1175	1%	1197	1%	0	0%	**2799**	**3%**
	White	36230	36%	29353	29%	20352	20%	2353	2%	**88289**	**88%**
	Total	**40919**	**41%**	**32617**	**33%**	**23944**	**24%**	**2353**	**2%**	**99834**	**100%**



Appendix

```
library(readxl)
library(survey)
library(srvyr)
library(tidyverse)
library(ggplot2)

# Step 1: Load Excel sheets
full_pop <- read_excel("G:\\Full Project Data.xlsx", sheet = "Full Target Population")
voter_guide_pop <- read_excel("G:\\Full Project Data.xlsx", sheet = "Voter Guide Population")
survey_pop <- read_excel("G:\\Full Project Data.xlsx ", sheet = "Survey Population")
survey_responses <- read_excel("G:\\Full Project Data.xlsx", sheet = "Survey Response Raw Data")

# Step 2: Data cleaning and preparation
survey_pop <- survey_pop %>%
  mutate(
    Age = 2025 - Birth_Year, # Assuming current year is 2025
    Age_Groups = case_when(
      Age >= 18 & Age <= 25 ~ "18-25",
      Age >= 26 & Age <= 35 ~ "26-35",
      Age >= 36 & Age <= 45 ~ "36-45",
      Age >= 46 & Age <= 55 ~ "46-55",
      Age >= 56 & Age <= 65 ~ "56-65",
      Age >= 66 ~ "66 and above",
      TRUE ~ NA_character_
    )
  )
colnames(full_pop) <- tolower(colnames(full_pop))
colnames(voter_guide_pop) <- tolower(colnames(voter_guide_pop))
colnames(survey_pop) <- tolower(colnames(survey_pop))
colnames(survey_responses) <- tolower(colnames(survey_responses))

# Step 3: Define population targets from Full Target Population
pop_targets <- list(
  age_groups = full_pop %>%
    count(age_groups) %>%
    rename(Freq = n),
  voter_political_party = full_pop %>%
    count(voter_political_party) %>%
    rename(Freq = n),
  sex = full_pop %>%
    count(sex) %>%
    rename(Freq = n), # Include if available in full_pop
```

```
  race_ethnicity = full_pop %>%
    count(race_ethnicity) %>%
    rename(Freq = n) # Include if available in full_pop
)

# Step 4: Nonresponse analysis
# Compare demographics of survey respondents (248) to survey population (4,150)
nonresponse_analysis <- function(var, survey_pop, survey_responses) {
  pop_dist <- survey_pop %>%
    count(.data[[var]]) %>%
    mutate(prop_pop = n / sum(n))
  resp_dist <- survey_responses %>%
    count(.data[[var]]) %>%
    mutate(prop_resp = n / sum(n))
  comparison <- left_join(pop_dist, resp_dist, by = var) %>%
    mutate(diff = prop_resp - prop_pop)
  return(comparison)
}

# Run nonresponse analysis for key variables
nonresponse_age <- nonresponse_analysis("age_groups", survey_pop, survey_responses)
nonresponse_party <- nonresponse_analysis("voter_political_party", survey_pop, survey_responses)
write.csv(nonresponse_age, "nonresponse_age.csv")
write.csv(nonresponse_party, "nonresponse_party.csv")

# Step 5: Create survey design object
# Assume no initial weights; we'll compute via raking
base_design <- svydesign(
  ids = ~survey_id, # Unique survey ID
  strata = NULL,
  weights = NULL,
  data = survey_responses
)

# Step 6: Raking to match Full Target Population
raked_design <- rake(
  base_design,
  sample.margins = list(~age_groups, ~voter_political_party),
  population.margins = list(pop_targets$age_groups, pop_targets$voter_political_party)
)

# Step 7: Generate crosstabs
crosstab_q1_age <- svyby(
```

```
  ~q1_helpfulness,
  ~age_groups,
  raked_design,
  svymean,
  na.rm = TRUE,
  vartype = c("se", "ci")
)
crosstab_q1_party <- svyby(
  ~q1_helpfulness,
  ~voter_political_party,
  raked_design,
  svymean,
  na.rm = TRUE,
  vartype = c("se", "ci")
)
write.csv(crosstab_q1_age, "crosstab_q1_age.csv")
write.csv(crosstab_q1_party, "crosstab_q1_party.csv")

# Step 8: Visualizations
# Bar plot for Q1_Helpfulness by Age Groups
ggplot(data = as.data.frame(crosstab_q1_age), aes(x = age_groups, y = q1_helpfulness)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  geom_errorbar(aes(ymin = q1_helpfulness - se, ymax = q1_helpfulness + se), width = 0.2) +
  labs(title = "Weighted Helpfulness (Q1) by Age Group", y = "Proportion", x = "Age Group") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
ggsave("q1_helpfulness_by_age.png")

# Bar plot for Q1_Helpfulness by Party
ggplot(data = as.data.frame(crosstab_q1_party), aes(x = voter_political_party, y = q1_helpfulness)) +
  geom_bar(stat = "identity", fill = "darkgreen") +
  geom_errorbar(aes(ymin = q1_helpfulness - se, ymax = q1_helpfulness + se), width = 0.2) +
  labs(title = "Weighted Helpfulness (Q1) by Party", y = "Proportion", x = "Party") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
ggsave("q1_helpfulness_by_party.png")
```

Similarly calculated weights for all other questions…..!!!!!!