

TASK 1: CONFIDENCE INTERVAL ESTIMATIONS

a) Given the information, construct a 95% confidence interval for “Approving” rate for each president in your dataset and comment your findings.

Column1	
Mean	40.38297872
Standard Error	0.262607063
Median	40
Mode	38
Standard Deviation	3.118286101
Sample Variance	9.723708207
Kurtosis	0.507595279
Skewness	0.685344911
Range	15
Minimum	34
Maximum	49
Sum	5694
Count	141
Confidence Level(95.0%)	0.519188271
lower limit	39.86379045
upper limit	40.90216699

Figure 1: Trump 95% confidence interval for Approving rate

Column1	
Mean	47.96889952
Standard Error	0.262607319
Median	47
Mode	46
Standard Deviation	5.36901933
Sample Variance	28.82636857
Kurtosis	1.684423712
Skewness	1.213709111
Range	27
Minimum	40
Maximum	67
Sum	20051
Count	418
Confidence Level(95.0%)	0.516199107
lower limit	47.45270041
upper limit	48.48509863

Figure 2:Obama 95% confidence interval for Approving rate

Column1	
Mean	51.34751773
Standard Error	0.945246058
Median	50.5
Mode	32
Standard Deviation	15.87338058
Sample Variance	251.9642109
Kurtosis	-0.558889046
Skewness	0.453670948
Range	65
Minimum	25
Maximum	90
Sum	14480
Count	282
Confidence Level(95.0%)	1.860662127
lower limit	49.4868556
upper limit	53.20817986

Figure 3:GWBush 95% confidence interval for Approving rate

Column1	
Mean	45.58333333
Standard Error	1.328228452
Median	42
Mode	41
Standard Deviation	6.50696394
Sample Variance	42.34057971
Kurtosis	-1.001554028
Skewness	0.84237061
Range	19
Minimum	38
Maximum	57
Sum	1094
Count	24
Confidence Level(95.0%)	2.747649896
lower limit	42.83568344
upper limit	48.33098323

Figure 4:Biden 95% confidence interval for Approving rate

Comments:

The 95% confidence interval for Approving rate for Biden – (42.83568344,48.33098323)

The 95% confidence interval for Approving rate for Trump – (39.86379045, 40.90216699)

The 95% confidence interval for Approving rate for Obama – (47.45270041, 48.48509863)

The 95% confidence interval for Approving rate for GWBush – (49.4868556, 53.20817986)

1) GW Bush has a higher mean approval rating than the other three presidents. It signifies that average approval rating is more for him.

2) Compared to the other three presidents, Biden values the standard error more. It implies that the estimate is more accurate.

3) GW Bush has a higher standard deviation number, which shows that the data are more evenly distributed.

4) The skewness of all presidents is positive; we can conclude that they are positively skewed distribution.

5) All presidents' confidence intervals, except for Biden, do not overlap, showing that their approval ratings differ from one another.

b) Which president has the narrowest/widest confidence intervals and what factors (i.e., sample size, sample variance, etc.) might contribute to this outcome. Obama has the narrowest confidence interval, while Biden has the biggest. The other variables, such as sample size, influence the confidence interval. The confidence interval is smaller when there are more samples, while the confidence interval is larger when there are fewer samples. Obama had more samples in this instance, which lowers the level of confidence.

The other components sample variance also impacts the confidence interval. Higher sample variance leads to largest confidence interval while lesser sample variance leads to narrowest confidence interval. The interval is also impacted by the confidence level. The interval is larger, the higher the confidence level. The confidence intervals are also influenced by this issue. The precision of confidence interval is lower for Biden approval rate because of difference between upper and lower limit is more than other presidents.

TASK 2: Hypothesis Testing

a) Use a one-sample t test to determine if “President Obama has a higher approval rating in the first 1000 days of his presidency than the last 1000 days”

In your answer, specify the competing hypothesis, the value of test statistics, and the p-value. You should clearly state the conclusion of your test (reject the null or fail to reject the null) at 5% significance level and the criteria you used to reach that conclusion.

TASK 2		
n for last 1000 days	✓	143
Mean for last 1000 days	✓	47.51748252
sample standard deviation for last 1000days	✓	4.30664921
Approval rate for last 1000 days		0.332290088
n for first 1000 days	✓	144
Mean fo first 1000 days	✓	50.02777778
sample standard deviation for first 1000 days	✓	6.841813488
Approval rate for first 1000 days		0.347415123
Pooled Standard Deviation		5.7065013
test static value		0.148178044
p value		0.441153521

Null Hypothesis: The approval rate of President Obama in his presidency for first 1000 days is less than or equal to approval rate in last 1000days.

Alternate Hypothesis: The approval rate of President Obama in his presidency for first 1000 days is greater than approval rate in last 1000days.

At 5% significance level, the p value is greater than 0.05. So, we do not reject the null hypothesis and Null hypothesis is true. There is no significance evidence to tell that approval rate for President Obama in first 1000 days is greater than in last 1000days.

b) Use a one-sample t test to determine if “President Trump has a same approval rating in the first 500 days of his presidency than the last 500 days”

In your answer, specify the competing hypothesis, the value of test statistics, and the p-value. You should clearly state the conclusion of your test (reject the null or fail to reject the null) at 1% significance level and the criteria you used to reach that conclusion.

Null Hypothesis: The approval rate of President Trump in his presidency for first 500 days is not equal to approval rate in last 1000days.

Alternate Hypothesis: The approval rate of President Trump in his presidency for first 500 days is equal to approval rate in last 1000days.

TASK 2				
n for last 500 days	✓	27		
Mean for last 500 days	✓	43.18518519		
sample standard deviation for last 500 days	✓	3.961711334		
Approval rate for last 500 days		1.599451303		
n for first 500 days	✓	72		
Mean for first 500 days	✓	38.875		
sample standard deviation for first 500 days	✓	2.207142694		
Approval rate for first 500 days		0.539930556		
Pooled Standard Deviation		2.778929548		
test static value		0.437633		
p- value		0.662625247		

At 1% significance level, the p value is greater than 0.01. So, we do not reject the null hypothesis and Null hypothesis is true. There is significance evidence to tell that approval rate for President Trump in first 1000 days is not equal to approval rate in last 1000days and there is no significance evidence to tell that approval rate for President Trump in first 1000 days is equal to approval rate in last 1000days

TASK 3: Statistical Inference Concerning Two populations

F	G	H	I	J
Count of Q3a		Column Labels ▼		
Row Labels ▼		Female	Male	Grand Total
Having a close relationship to Germany		241	254	495
Having a close relationship to Russia		90	100	190
Grand Total		331	354	685

		Female	Male	Total
Having a close relationship to Germany	Observed	241	254	495
	Expected	239.19	255.81	495.00
Having a close relationship to Russia	Observed	90	100	190
	Expected	91.81	98.19	190.00
Total	Observed	331	354	685
	Expected	331.00	354.00	685.00

.10 chi-square
1 df
.7572 p-value
.012 Cramér's V

a) Fill out the following 2 by 2 tables by including both the expected and observed frequencies for each cell.

Observed Frequencies	Female	Male	Grand Total
Having a close relationship to Germany	241	254	495
Having a close relationship to Russia	90	100	190
Grand Total	331	354	685

Expected Frequencies	Female	Male	Grand Total
Having a close relationship to Germany	239.19	255.81	495.00
Having a close relationship to Russia	91.81	98.19	190.00

Grand Total	331.00	354.00	685.00
--------------------	--------	--------	--------

b) State the Null and the Alternate Hypothesis:

Null Hypothesis: There is no significant connection between genders in sex column and having a close relationship to Germany or Russia in Q3a column.

Alternate Hypothesis: There is significant connection between genders in sex column and having a close relationship to Germany or Russia in Q3a column.

c) Conduct a Chi-squared test and fill out the following table from your excel output:

	Chi-square test statistics
Degrees of freedom (df)	1
Value	0.10
P-value	0.7572

Based on results, we cannot reject the null hypothesis at 5% significance level because the P-value is greater than significance value. Now we can say that there is no significant connection between genders in sex column and having a close relationship to Germany or Russia in Q3a column. In other words, null hypothesis is true.

d) Calculate the Cramer's V value and interpret it.

The Cramer's V value ranges from zero to one. From this range, we can infer that there is a weak to strong correlation between the variables, ranging from 0 to 1. So, here the value is **0.012**. It is very close to zero saying that the null hypothesis is true, and we cannot reject the null hypothesis.

Count of partyIn	Column Labels		
Row Labels	Republican	Democratic	Grand Total
Likely	194	221	415
Unlikely	131	139	270
Grand Total	325	360	685

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

.017 Cramér's V

By responding to the questions listed below, I'd rate the possibility of the current conflict. We must first state the null and alternative hypotheses. They help us clearly grasp what we need to look for. Calculate specified values, such as observed values, expected values, p values, etc. in

accordance with the hypotheses. These criteria can be used to evaluate how people feel about the US-China Cold War. We can determine whether to reject the null hypothesis at a specific significance level. With the aid of opinions from democratic and republican people, we may use this to estimate the likelihood of conflict between China and the US.

a) Generate a two-by-two table in Excel, put Q4 answers (Likely, Unlikely) in rows and political leanings (Republican, Democratic) in columns. In your table, make sure both the frequency counts and expected values are listed. Insert your tables below.

Observed	Republican	Democratic	Total
Likely	194	221	415
Unlikely	131	139	270
Total	325	360	685

Expected	Republican	Democratic	Total
Likely	196.90	218.10	415.00
Unlikely	128.10	141.90	270.00
Total	325.00	360.00	685.00

b) State the Null and the Alternate Hypothesis:

Null Hypothesis: There is no significant connection between the people with two different political leanings and their views on likelihood of cold war happening between China and US.

Alternate Hypothesis: There is significant connection between the people with two different political leanings and their views on likelihood of cold war happening between China and US.

c) Conduct a Chi-squared test and fill out the following table from your Excel Output:

	Chi-square test statistics
Degrees of freedom (df)	1
Value	0.21
P-value	0.6500

d) Based on your results, would you reject the null hypothesis at 5% significance level?

Based on results, we cannot reject the null hypothesis at 5% significance level because the P-value is greater than significance value. Now we can say that there is no significant connection

between the people with two different political leanings and their views on likelihood of cold war happening between China and US. In other words, null hypothesis is true.

TASK 5: Statistical Inference Concerning Two populations

Count of party in	Column Labels			
Row Labels	Republican	Democratic	Grand Total	
High-educated	182	210	392	
Grand Total	182	210	392	
$\hat{p}_{High-educated}^{Rep}$	0.464285714			
$\hat{p}_{High-educated}^{Dem}$	0.535714286			
Point estimate of our parameter of interest	-0.071428571			
standard error for the point estimate	0.050507627			
z	-1.644853627			
lower limit	-0.154506225			
upper limit	0.011649082			

a) Calculate the point estimate of our parameter of interest, $\hat{p}_{High-educated}^{Rep} - \hat{p}_{High-educated}^{Dem}$

The point estimate of our parameter of interest is -0.71428571

b) Assuming that the population parameter is normally distributed and all conditions for confidence interval calculations for proportion differences met, calculate the appropriate standard error for the point estimate, $\hat{p}_{High-educated}^{Rep} - \hat{p}_{High-educated}^{Dem}$

Standard error for the point estimate is 0.050507627.

c) Construct a 90% confidence interval for proportion differences:

By constructing a confidence interval in excel shown above, we got below confidence interval values for upper limit and lower limit.

The 90% confidence interval is (-0.154506225, 0.011649082)

d) Based on the 90% confidence interval, what can you say about our research question? Do you have statistically significant evidence to say that the level of education differs across political leanings: proportion of Republican-leaning Americans who are consider as highly educated differ from their democratic leaning counterpart? Please explain.

No, we cannot decide the research question based on 90% confidence interval. First point, the zero can include between width of interval. If it is zero, there is no significant difference between Republican highly educated people and Democratic highly educated people. So, we there is no significant evidence to say that there is difference between Republican leaning American and Democratic leaning Americans who are considered as highly educated.