Mutaal Data Brief #2

Mutaal Tahir

2022-11-28

Research Topic and Question

This project studies the difference in opinions towards government response and restrictions due to the COVID-19 pandemic among people from rural and urban Alberta. To address this topic, I will analyze the September 2021 Alberta Viewpoint Data set. The motivation behind choosing this data set is that it marked over a year of COVID-19 restrictions being in place in Alberta and I aim to study how they have felt so far with the government's response to the pandemic particularly from a rural/urban perspective. Do the opinions of rural residents differ from their counterparts in the urban areas or do they feel similar towards the government's response and restrictions? I will look into the responses based on both a provincial and federal government level.

Prior research shows that a rural urban divide is present in Alberta (Banack 2021, Epp 2019). This divide is linked to factors and has caused rural residents to feel overlooked, looked down upon and treated unfairly (Banack, 2021). Epp (2019) argues that rural sensibility is hard to find in the government caucus, which reflects on Alberta becoming an increasingly urban province compared to a rural one. An example of this difference among rural and urban residents in Canada is the carbon tax where 48% of rural residents believe it should be repealed compared to 34.3% or urban residents (Sean Speer et al., 2021). One reason for this is the fact that the carbon tax affects rural residents disproportionately as they have higher transportation costs and limited public transit (Epp 2019). An example Epp (2019) gives is the 2016 Carbon Tax which did not consider places like Foremost and Tulliby Lake where driving long distances is a daily activity and so a higher cost has to be paid by local residents. Speer et al. (2021) also shows that rural residents have a stronger belief in emphasis on tradition compared to urban residents. This conclusion has also been found in research done in the United states where rural residents were more likely to be traditional in outlook and values (Sean Speer et al., 2021). This may also explain why in Canada the Conservatives have had a strong hold in rural areas during various times including the present era since 2004 (Armstrong et al., 2021).

COVID-19 may have brought more issues that may have fueled this divide. Among these issues may include shining a light on Alberta's rural healthcare system, which has already had issues including the lack of access for osteoarthritis care despite the high prevalence (Liu et al., 2022). In 2013, the Canadian Medical Association reported that 21% of the Canadian population lives in rural areas, but only 9.4% of general practitioners and 3% of specialists do (CMA, 2013). COVID-19 causes many infected individuals to be hospitalized, which could prove to be a problem for the already understaffed rural healthcare system. Torrie et al. (2021) highlights how there have been vaccine distribution issues in rural Alberta in the past giving an example of the 2009 HINI influenza pandemic. Although these issues are related to healthcare and my project is studying opinions on government policies, it is important to highlight these issues as they may affect the opinions of Albertans living on the rural and urban side regarding government policies since they may hold the government responsible for these inadequacies.

Data and Methods

The 2021 Alberta Viewpoint Survey contains data from 1204 respondents. Data was collected through a roughly 15-minute voluntary survey. Consent of respondents was taken before the survey began. Respondents included individuals from various backgrounds, age, demographics and genders. The survey had 8 sections namely:

1. Demographics
2. Electoral Politics
3. Campaign Finances and Local Elections
4. Dialogue and Polarization
5. Federalism
6. The COVID-19 Pandemic
7. Economic Perceptions
8. Identity and Discrimination

Missing information in the responses and was deleted using list wise deletion in order to have clean data to analyze.

I will be analyzing data using bivariate analyses with cross tables, Cramer's V test, Chi square test and visualizations. Using cross tables I will set the conditional argument as the zone i.e. rural or urban. This will help me finding the probabilities of events occurring given the respondents location. Since all my outcome variables are ordinal and predictor variables are nominal, I will be using a Cramer's V test to test the association between my variables as that supports ordinal and nominal variables being used together. The association I want to test is whether area zone affects viewpoints of our respondents or not. Furthermore, I will also use a chi square test to check whether the relationships between viewpoints and area zones are statistically significant or not.

```
d1 <- read.csv("ViewpointAB_Sept2021_Survey_Data.csv",header=TRUE)
d2 <- subset(d1,select=c(Q4,Q61r1,Q61r2,Q67,Q68,Q70))
d3 <- subset(d2,complete.cases(d2),drop=TRUE)
dim(d1)
```

```
## [1] 1204  343
```

Analysis

For ease of reading and simplicity I will divide each variable into its own category and describe it.

Area

Areas are divided into three categories: urban, suburban and rural. Urban has been coded into '1', Suburban '2' and rural '3'. I have recoded them into their respective names so that the viewer does not have to look back and forth for what each number represents. I also counted suburban residents as urban residents in this project due to their close proximity with urban areas.

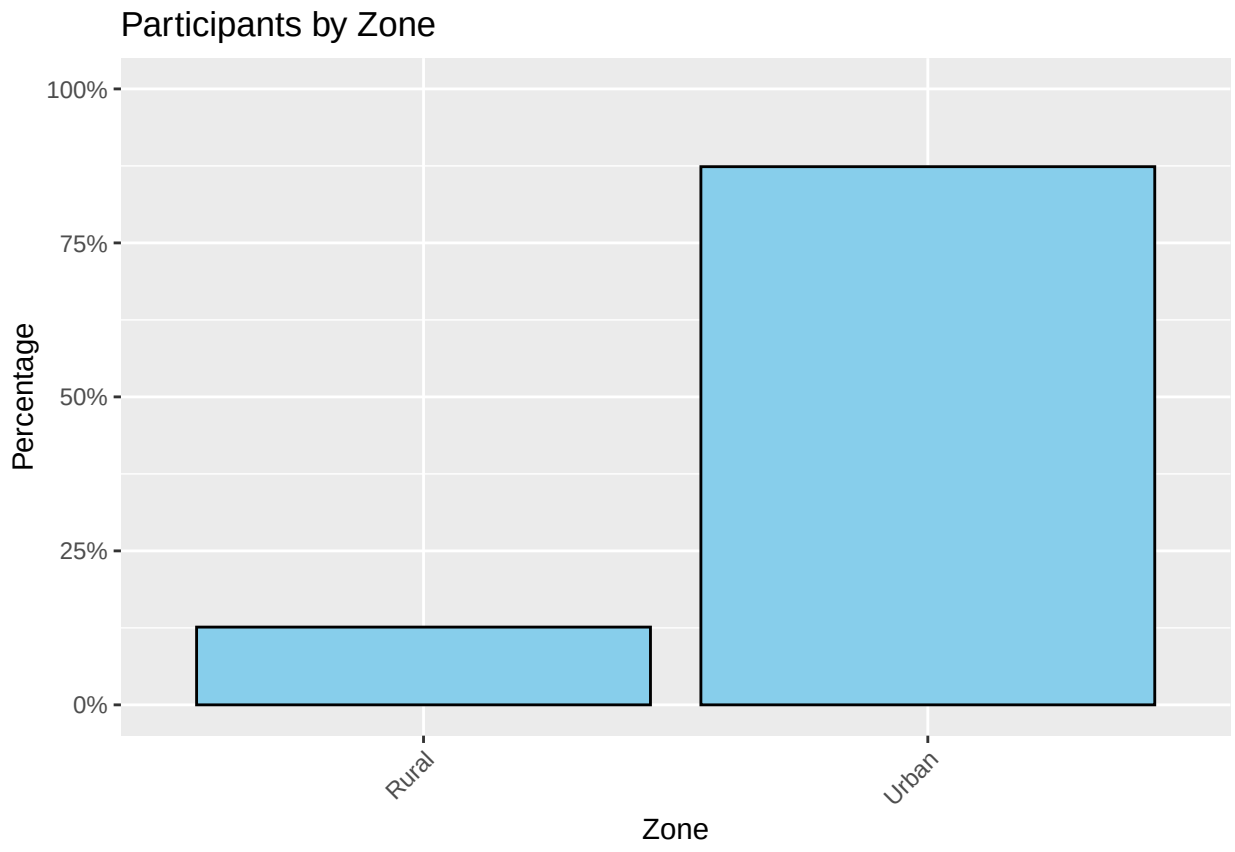
```
d3$zone=recode(d3$Q4,`1`="Urban",`2`="Urban",`3`="Rural")
table(d3$zone)
```

```
##
## Rural Urban
##   152  1052
```

```
round(prop.table(table(d3$zone)) * 100, 2)
```

```
##
## Rural Urban
## 12.62 87.38
```

```
ggplot(data = d3, aes(x = zone, y = ..count.. / sum(..count..))) + geom_bar(fill = "sky blue", color =
labs(x = "Zone",
y = "Percentage",
title = "Participants by Zone") +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Data shows that 87.38% (1052) of our respondents are from urban areas whereas only 12.62% (152) are from rural areas.

Performance of Government (Federal)

This is an outcome variable I will need to predict how well do rural vs urban Albertans think the Federal government handled the pandemic. Values are in numbers and were recoded as: '1' Very bad, '2' Bad, '3' Good, '4' Very Good.

In order to measure conditional probability I would need to use cross tables. I want to measure what is the probability that someone will rate federal performance on a certain level given their area status.

```
d3$trust_fg=recode(d3$Q61r1,`1`="Very bad",`2`="Bad",`3`="Good",`4`="Very Good")
d3$trust_fg <- factor(d3$trust_fg,c("Very bad","Bad","Good","Very Good"))
prop.table(table(d3$zone, d3$trust_fg), margin = 1)
```

```
##
##           Very bad        Bad       Good  Very Good
##   Rural 0.31578947 0.28289474 0.38157895 0.01973684
##   Urban 0.21482890 0.23954373 0.46007605 0.08555133
```

Given that the residents were rural, 31.58% thought the federal government did a very bad job, 28.29% thought the federal government did a bad job, 38.16% thought the federal government did a good job, 19.73% thought the federal government did a very good job.

Given that the residents were urban, 21.48% thought the federal government did a very bad job, 23.95% thought the federal government did a bad job, 46.01% thought the federal government did a good job, 8.56% thought the federal government did a very good job.

```
library(desc)
```

```
##
## Attaching package: 'desc'

## The following object is masked from 'package:dplyr':
##
##     desc
```

```
CramerV(d3$zone,d3$trust_fg)
```

```
## [1] 0.1161136
```

When using a Cramer's V test for an association between area zone and performance of federal we get a value of 0.116 which indicates a weak association.

I will now use a Chi square test to see if there is an association between trust in the federal government and the area zone.

```
chisq.test(d3$trust_fg,d3$zone, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  d3$trust_fg and d3$zone
## X-squared = 16.233, df = NA, p-value = 0.0009995
```

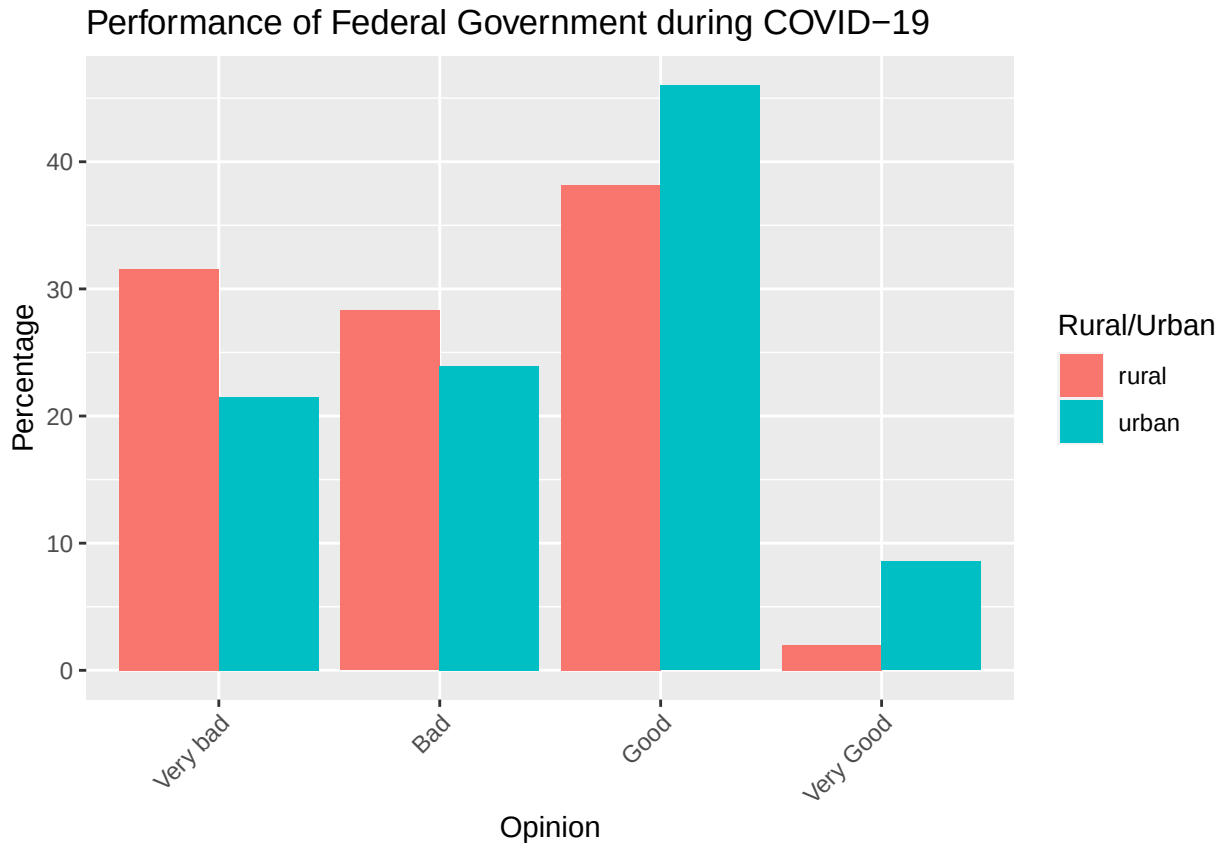
Since we get a p-value of <0.05 we reject the null hypothesis meaning that there is an statistically significant relationship between area zone and trust in the federal government.

```
data1 <- data.frame(values = as.vector(round(prop.table(table(d3$zone, d3$trust_fg), margin = 1)*100,2))
zone = rep(c("rural","urban"),4)
, trust.fg = c(rep("Very bad",2), rep("Bad",2), rep("Good",2), rep("Very Good",2)))

data1$zone <- factor(data1$zone, levels = c("rural", "urban"))

data1$trust.fg <- factor(data1$trust.fg, levels = c("Very bad","Bad","Good","Very Good"))

ggplot(data = data1, aes(y = values, x = trust.fg, fill = zone)) + geom_bar(position = "dodge", stat="ic
labs(x = "Opinion",
y = "Percentage",
title = "Performance of Federal Government during COVID-19", fill = "Rural/Urban") +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Performance of Government (Provincial)

This is an outcome variable I will need to predict how well do rural vs urban Albertans think the provincial government handled the pandemic. Values are in numbers and were recoded as: '1' Very bad, '2' Bad, '3' Good, '4' Very Good.

In order to measure conditional probability I would need to use cross tables. I want to measure what is the probability that someone will rate federal performance on a certain level given their area status.

```
d3$trust_pg=recode(d3$Q61r2,`1`="Very bad",`2`="Bad",`3`="Good",`4`="Very Good")
d3$trust_pg <- factor(d3$trust_pg,c("Very bad","Bad","Good","Very Good"))
prop.table(table(d3$zone, d3$trust_pg), margin = 1)
```

```
##
##            Very bad         Bad        Good   Very Good
##   Rural 0.506578947 0.322368421 0.164473684 0.006578947
##   Urban 0.547528517 0.267110266 0.154942966 0.030418251
```

Given that the residents were rural 50.66% thought the provincial government did a very bad job, 32.23% thought the provincial government did a bad job, 16.44% thought the provincial government did a good job, 0.65% thought the provincial government did a very good job.

Given that the residents were urban 54.75% thought the provincial government did a very bad job, 26.71% thought the provincial government did a bad job, 15.49% thought the provincial government did a good job, 3.04% thought the provincial government did a very good job.

```
library(desc)
CramerV(d3$zone,d3$trust_pg)
```

```
## [1] 0.06262041
```

When using a Cramer's V test for an association between area zone and performance of provincial government we get a value of 0.063 which indicates a weak association.

I will now use a Chi square test to see if there is an association between trust in the provincial government and the area zone.

```
chisq.test(d3$trust_pg,d3$zone, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  d3$trust_pg and d3$zone
## X-squared = 4.7213, df = NA, p-value = 0.1724
```

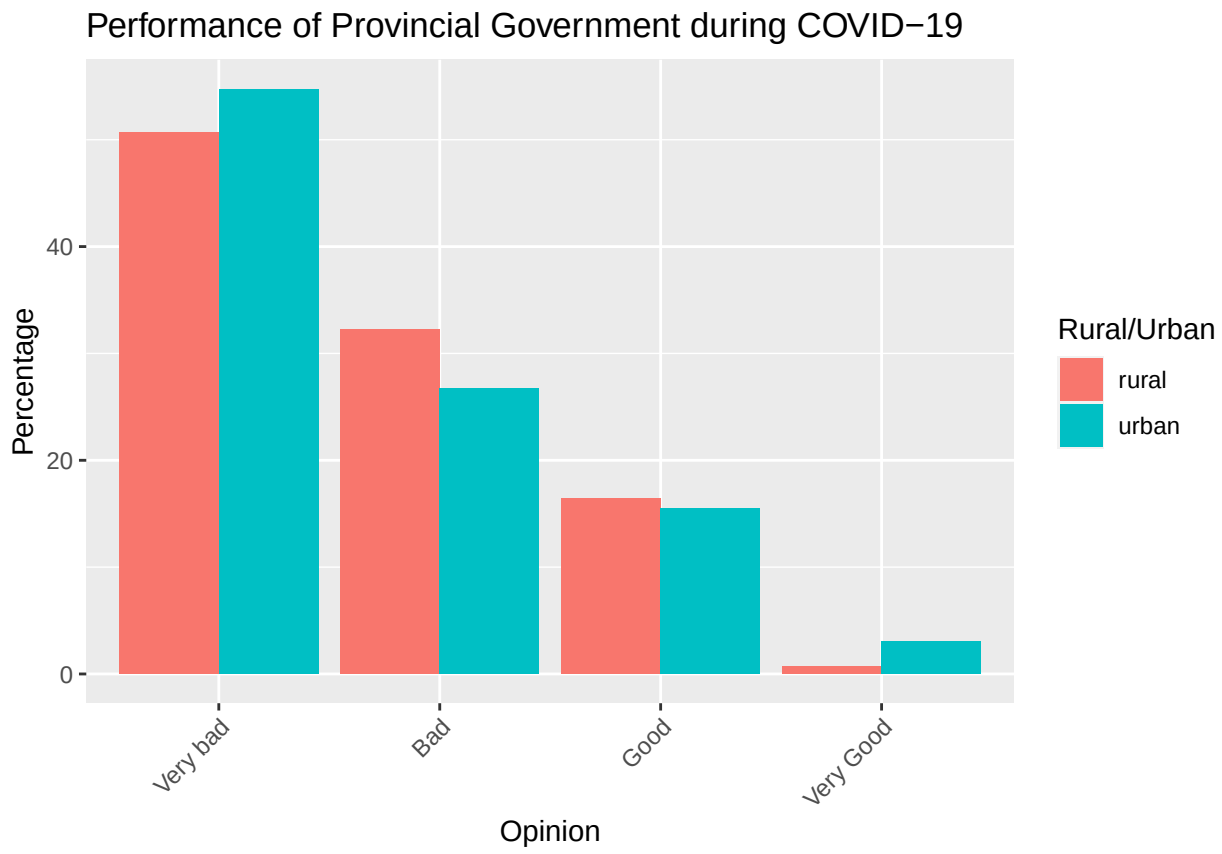
Since we get a p-value of >0.05 we fail to reject the null hypothesis meaning that there isn't a statistically significant relationship between area zone and trust in the provincial government.

```
data2 <- data.frame(values = as.vector(round(prop.table(table(d3$zone, d3$trust_pg), margin = 1)*100,2))
zone = rep(c("rural","urban"),4)
, trust.pg = c(rep("Very bad",2), rep("Bad",2), rep("Good",2), rep("Very Good",2)))

data2$zone <- factor(data2$zone, levels = c("rural", "urban"))

data2$trust.pg <- factor(data2$trust.pg, levels = c("Very bad","Bad","Good","Very Good"))

ggplot(data = data2, aes(y = values, x = trust.pg, fill = zone)) + geom_bar(position = "dodge", stat="i
labs(x = "Opinion",
y = "Percentage",
title = "Performance of Provincial Government during COVID-19", fill = "Rural/Urban") +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Support of businesses requiring proof of vaccination (Q67)

This is an outcome variable I will need to predict how well do rural vs urban Albertans support the government law allowing businesses to stay open provided they require proof of vaccination of negative COVID-19 test results. Values are in numbers and were recoded as: '1' Strongly support, '2' Somewhat support, '3' Neither support or oppose, '4' Somewhat oppose, '5' Strongly oppose.

```
d3$business_vac=recode(d3$Q67,`1`="Strongly support",`2`="Somewhat support",`3`="Neither support or oppo
d3$business_vac <- factor(d3$business_vac,c("Strongly support","Somewhat support","Neither support or o
prop.table(table(d3$zone, d3$business_vac), margin = 1)
```

```
##
##         Strongly support Somewhat support Neither support or oppose
##   Rural       0.46052632       0.18421053                0.09868421
##   Urban       0.49714829       0.23574144                0.09790875
##
##         Somewhat oppose Strongly oppose
##   Rural      0.05921053      0.19736842
##   Urban      0.04277567      0.12642586
```

Given that the residents were rural 46.05% strongly supported proof of vaccinations, 18.42% somewhat supported proof of vaccinations, 9.87% neither supported nor opposed proof of vaccinations, 5.92% somewhat opposed proof of vaccinations and 19.74% strongly opposed proof of vaccinations.

Given that the residents were urban 49.71% strongly supported proof of vaccinations, 23.57% somewhat supported proof of vaccinations, 9.79% neither supported nor opposed proof of vaccinations, 4.28% somewhat opposed proof of vaccinations and 12.64% strongly opposed proof of vaccinations.

```
library(desc)
CramerV(d3$zone,d3$business_vac)
```

```
## [1] 0.07964836
```

When using a Cramer's V test for an association between area zones and support for proof of vaccination by businesses we get a value of 0.079 which indicates a weak association.

I will now use a Chi square test to see if there is an association between support for requiring proof of vaccination and the area zone.

```
chisq.test(d3$business_vac,d3$zone, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  d3$business_vac and d3$zone
## X-squared = 7.638, df = NA, p-value = 0.1049
```

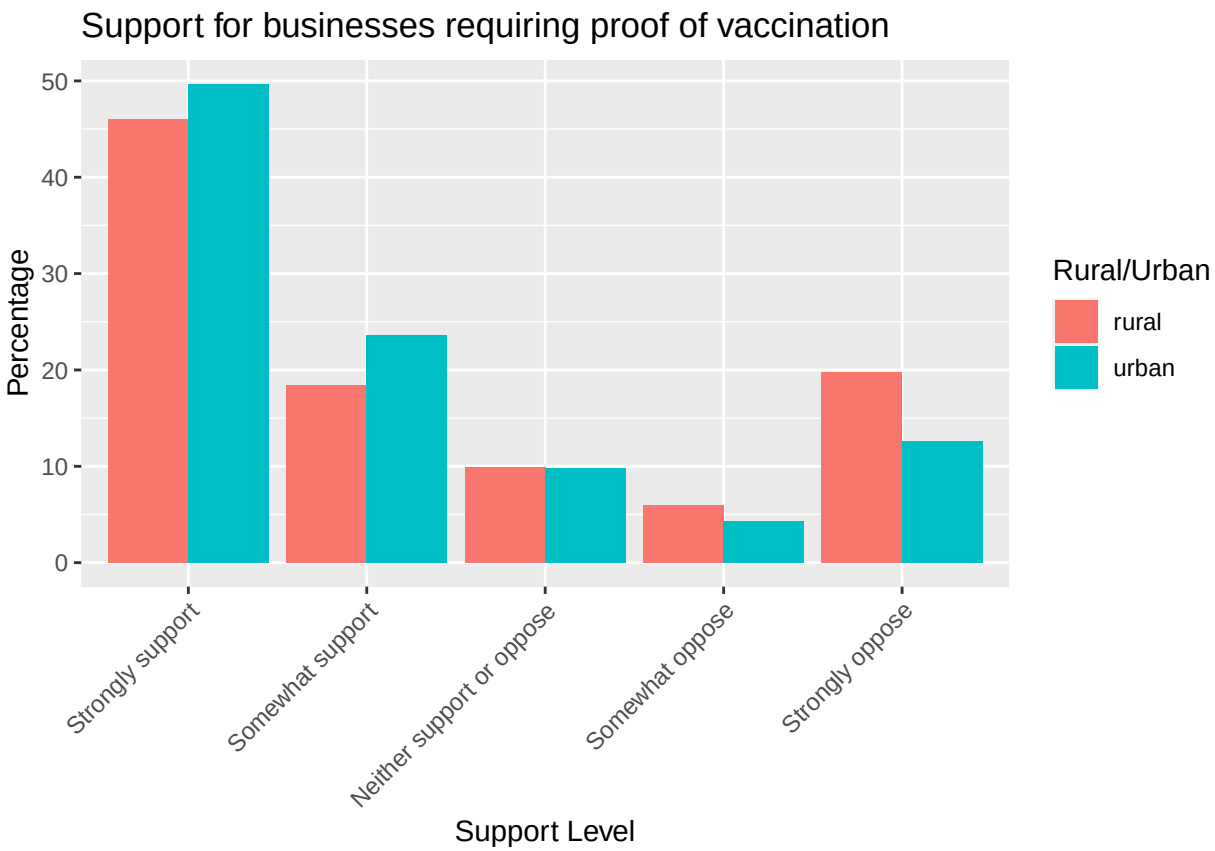
Since we get a p-value of >0.05 we fail to reject the null hypothesis meaning that there isn't a statistically significant relationship between area zone and support for businesses requiring proof of vaccination.

```
data3 <- data.frame(values = as.vector(round(prop.table(table(d3$zone, d3$business_vac), margin = 1)*100
zone = rep(c("rural","urban"),5)
, business.vac = c(rep("Strongly support",2), rep("Somewhat support",2), rep("Neither support or oppose

data3$zone <- factor(data3$zone, levels = c("rural", "urban"))

data3$business.vac <- factor(data3$business.vac, levels = c("Strongly support","Somewhat support","Neitl

ggplot(data = data3, aes(y = values, x = business.vac, fill = zone)) + geom_bar(position = "dodge", sta
labs(x = "Support Level",
y = "Percentage",
title = "Support for businesses requiring proof of vaccination", fill = "Rural/Urban") +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Support for vaccine passports (Q68)

This is an outcome variable I will need to predict how well rural vs urban Albertans support vaccine passports. Values are in numbers and were recoded as: '1' Strongly support, '2' Somewhat support, '3' Neither support or oppose, '4' Somewhat oppose, '5' Strongly oppose.

```
d3$vac_pass=recode(d3$Q68,`1`="Strongly support",`2`="Somewhat support",`3`="Neither support or oppose"
d3$vac_pass <- factor(d3$vac_pass,c("Strongly support","Somewhat support","Neither support or oppose","S
prop.table(table(d3$zone, d3$vac_pass), margin = 1)
```

```
##
##         Strongly support Somewhat support Neither support or oppose
##   Rural       0.55263158       0.09868421                0.07236842
##   Urban       0.62737643       0.14448669                0.07319392
##
##         Somewhat oppose Strongly oppose
##   Rural      0.05263158      0.22368421
##   Urban      0.02946768      0.12547529
```

Given that the residents were rural 55.26% strongly supported vaccine passports, 9.87% somewhat supported vaccine passports, 7.24% neither supported nor opposed vaccine passports, 5.26% somewhat opposed vaccine passports and 22.37% strongly opposed vaccine passports.

Given that the residents were urban 62.74% strongly supported vaccine passports, 14.44% somewhat supported vaccine passports, 7.32% neither supported nor opposed vaccine passports, 2.95% somewhat opposed vaccine passports and 12.55% strongly opposed vaccine passports.

```
library(desc)
CramerV(d3$zone,d3$vac_pass)
```

```
## [1] 0.1105013
```

When using a Cramer's V test for an association between area zones and support for vaccine passports we get a value of 0.111 which indicates a weak association. This is also reflected in our cross table which shows that results remained reasonably consistent despite change in area zone.

I will now use a Chi square test to see if there is an association between support for vaccine passports and the area zone.

```
chisq.test(d3$vac_pass,d3$zone, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  d3$vac_pass and d3$zone
## X-squared = 14.701, df = NA, p-value = 0.005997
```

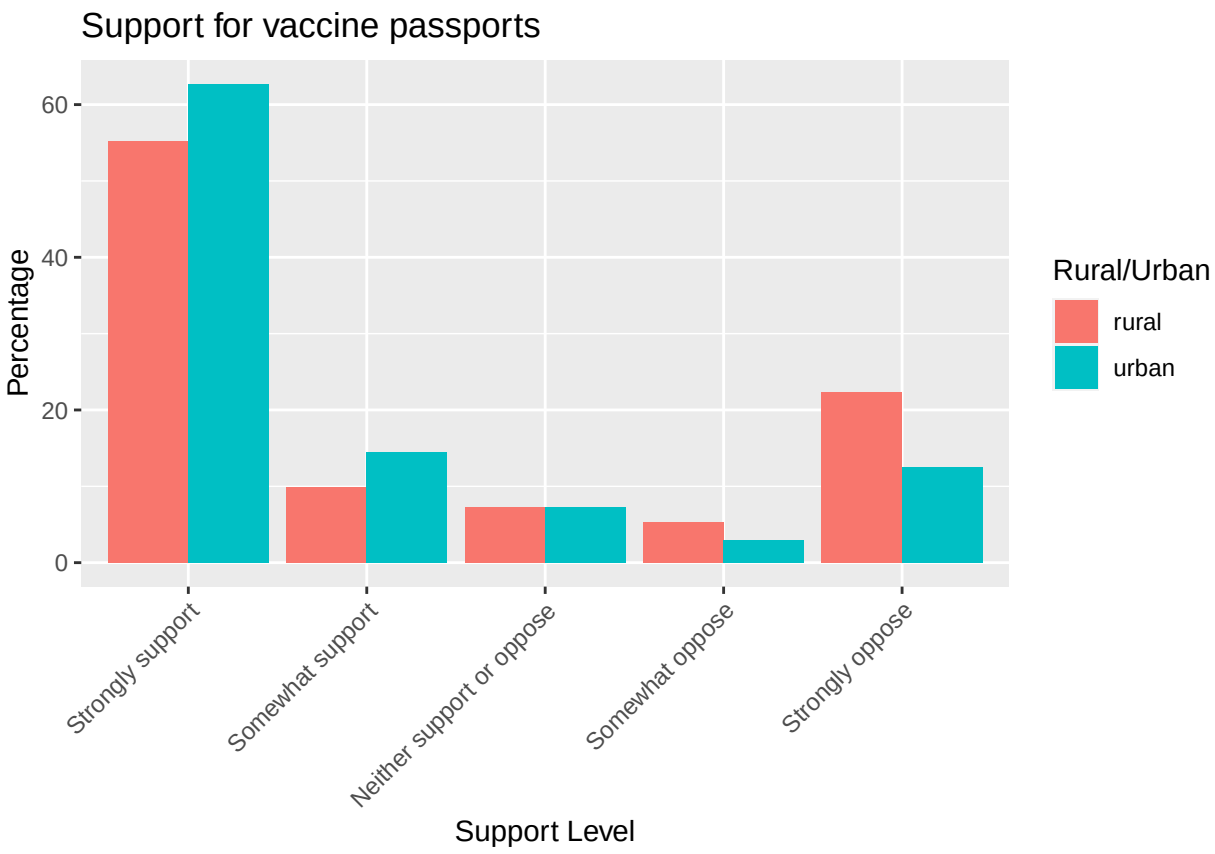
Since we get a p-value of <0.05 we reject the null hypothesis meaning that there is a statistically significant relationship between support for vaccine passports and the area zone.

```
data4 <- data.frame(values = as.vector(round(prop.table(table(d3$zone, d3$vac_pass), margin = 1)*100,2))
zone = rep(c("rural","urban"),5)
, vac.pass = c(rep("Strongly support",2), rep("Somewhat support",2), rep("Neither support or oppose",2)

data4$zone <- factor(data4$zone, levels = c("rural", "urban"))

data4$vac.pass <- factor(data4$vac.pass, levels = c("Strongly support","Somewhat support","Neither suppo

ggplot(data = data4, aes(y = values, x = vac.pass, fill = zone)) + geom_bar(position = "dodge", stat="i
labs(x = "Support Level",
y = "Percentage",
title = "Support for vaccine passports", fill = "Rural/Urban") +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Feelings on current restrictions (Q70)

This is an outcome variable I will need to predict how well rural vs urban Albertans feel about current restrictions. Values are in numbers and were recoded as: '1' Too harsh, '2' Too lenient, '3' Just about right.

```
d3$feel_cr=recode(d3$Q70,`1`="Too harsh",`2`="Too lenient",`3`="Just about right")
d3$feel_cr <- factor(d3$feel_cr,c("Too harsh","Too lenient","Just about right"))
prop.table(table(d3$zone, d3$feel_cr), margin = 1)
```

```
##
##         Too harsh Too lenient Just about right
##   Rural 0.2302632   0.5394737        0.2302632
##   Urban 0.1482890   0.5760456        0.2756654
```

Given that the residents were rural 23.03% thought current restrictions were too harsh, 53.95% thought they were too lenient, 23.03% thought they were just about right.

Given that the residents were urban 14.83% thought current restrictions were too harsh, 57.60% thought they were too lenient, 27.57% thought they were just about right.

```
library(desc)
CramerV(d3$zone,d3$feel_cr)
```

```
## [1] 0.07598098
```

When using a Cramer's V test for an association between area zone and feelings on current restrictions we get a value of 0.076 which indicates a weak association. This is also reflected in our cross table which shows that results remained reasonably consistent despite change in area zone.

I will now use a Chi square test to see if there is an association between beliefs on current restrictions and the area zone.

```
chisq.test(d3$feel_cr,d3$zone, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  d3$feel_cr and d3$zone
## X-squared = 6.9508, df = NA, p-value = 0.02749
```

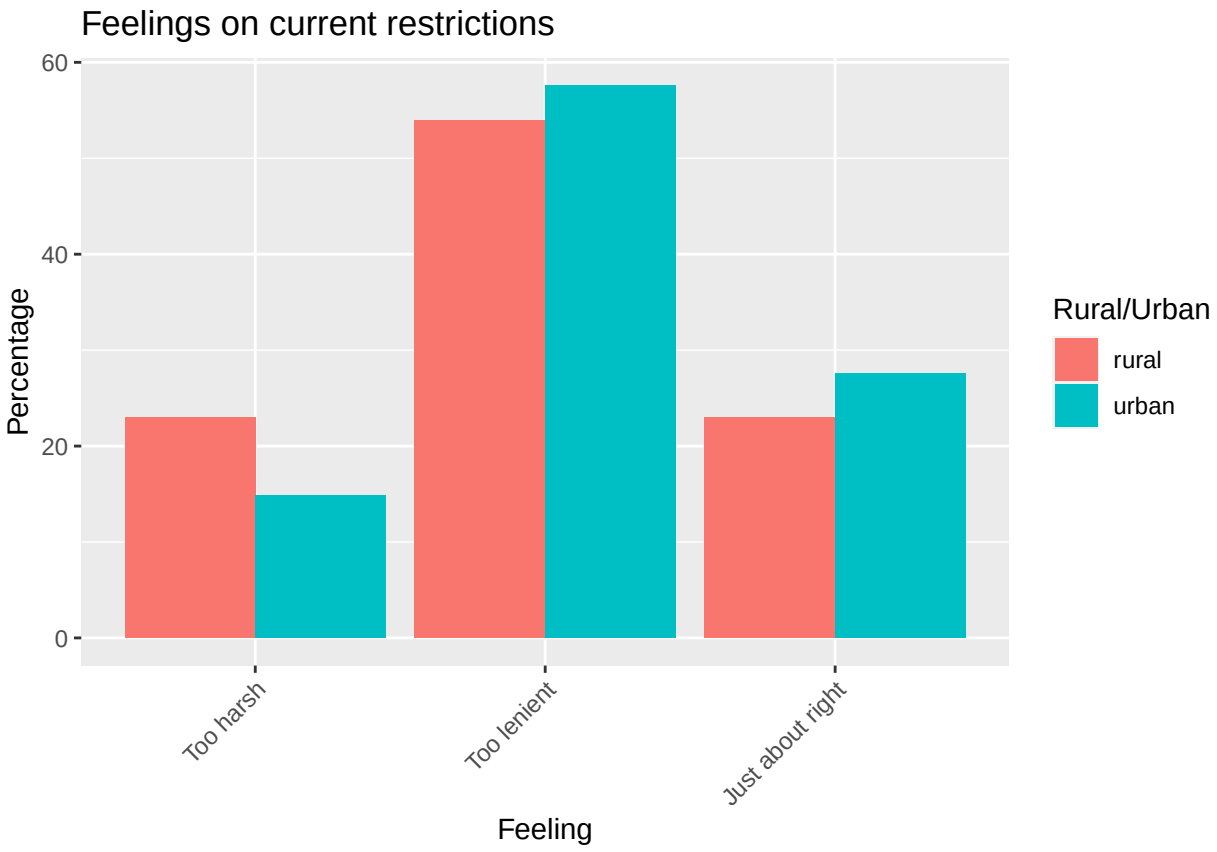
Since we get a p-value of <0.05 we reject the null hypothesis meaning that there is a statistically significant relationship between beliefs on current restrictions and the area zone.

```
data5 <- data.frame(values = as.vector(round(prop.table(table(d3$zone, d3$feel_cr), margin = 1)*100,2))
zone = rep(c("rural","urban"),3)
, feel.cr = c(rep("Too harsh",2), rep("Too lenient",2), rep("Just about right",2)))

data5$zone <- factor(data5$zone, levels = c("rural", "urban"))

data5$feel.cr <- factor(data5$feel.cr, levels = c("Too harsh","Too lenient","Just about right"))

ggplot(data = data5, aes(y = values, x = feel.cr, fill = zone)) + geom_bar(position = "dodge", stat="ide
labs(x = "Feeling",
y = "Percentage",
title = "Feelings on current restrictions", fill = "Rural/Urban") +
theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Results and Limitations

After doing the analyses using cross tables, Cramer's V and bar plots I found that there is a weak association between living zone and viewpoint regarding the government's COVID-19 policy. However, a chi square test revealed that in some cases a significant relationship may exist between zone area and the variable of measure. Statistically significant relationships existed in the performance of the federal government, vaccine passports and feelings on current restrictions. It was observed that in some cases one side was much more likely to support or oppose some measure than the other. One example of this includes vaccine passports where 22.37% of rural residents strongly opposed vaccines passports compared to 12.55% of urban residents. Another example is the performance of the federal government where 19.73% rural residents thought the federal government did a very good job compared to only 8.56% of urban residents. Moreover, 23.03% of rural residents found current restrictions too harsh compared to 14.82% urban residents. When it came to performance of the provincial government and businesses requiring proof of vaccine or negative COVID-19 there were no significant relationships found among area zones.

Limitations for this analysis included the fact that the sample size is very small compared to the general population of Alberta at the present moment which is about 4.5 million (*Population Statistics*, n.d.) whereas our sample contains only 1204 responses. Another limitation includes the fact that urban residents made 87.38% (1052) of our respondents whereas only 12.62% (152) were from rural areas. This leads to an even smaller rural representation for our study and so we would need to be very careful before making any generalization about rural viewpoints. Moreover, In a case where this data would be used to study any topic for the whole of Alberta urban opinions may end up dominating rural opinions which is why it is necessary to divide the data into groups. Nevertheless, we were able to gain various useful insights from this study and we can use this to gain a useful understanding of the current viewpoints of various Alberta residents up til September of 2021.

References

Banack, C. (2020, October 26). Ethnography and Political Opinion: Identity, Alienation and Anti-establishmentarianism in Rural Alberta. Canadian Journal of Political Science, 54(1), 1–22. https://doi.org/10.1017/s0008423920000694

Epp, Roger. 2019. "The End of Exceptionalism: Post-rural Politics in Alberta." In Orange Chinook: Politics in the New Alberta, ed. Duane Bratt, Keith Brownsey, Richard Sutherland and David Taras. Calgary: University of Calgary Press.

Liu, X.; Seidel, J.E.; McDonald, T.; Patel, A.B.; Waters, N.; Bertazzon, S.; Shahid, R.; Marshall, D.A. Rural–Urban Disparities in Realized Spatial Access to General Practitioners, Orthopedic Surgeons, and Physiotherapists among People with Osteoarthritis in Alberta, Canada. Int. J. Environ. Res. Public Health**2022**,19,7706. https://doi.org/10.3390/ijerph19137706

Canadian Medical Association. CMA Position Statement: Ensuring Equitable Access to Care: Strategies for Governments, Health System Planners, and the Medical Profession; Canadian Medical Association: Ottawa, ON, Canada, 2013.

Torrie, C., Yanicki, S., Sedgwick, M., & Howard, L. (2021, February 1). Social justice in pandemic immunization policy: We're all in this together. *Nursing Ethics*, *28*(6), 924–934. https://doi.org/10.1177/0969733020983395

Armstrong, D. A., Lucas, J., & Taylor, Z. (2021). The Urban-Rural Divide in Canadian Federal Elections, 1896–2019. Canadian Journal of Political Science, 55(1), 84–106. https://doi.org/10.1017/s000842392100079

Sean Speer, Peter Loewen, & Stephanie Bertolo. (2021). Perceptions and Polarization: Measuring the Perception Gap Between Urban and Rural Canadians. In *https://ppforum.ca/publications/* (ISBN: 978-1-77452-047-5). Public Policy Forum. https://ppforum.ca/publications/fault-lines-and-common-ground/

Population statistics. (n.d.). Alberta.ca. https://www.alberta.ca/population-statistics.aspx