Please read the article attached and then read each summary below of that article. Keep in mind that the way text summarizers work is they select what they consider the 3 or 4 most representative sentences in the document as the summary. Answer the questions following the summaries.

# Summary B (relStruct)

Although the prevalence of current cigarette smoking among US adults has decreased over the past several decades to 13.7% in 2018, this decrease has not been as pronounced in rural areas, such as rural Appalachia (as high as 33% in some counties). To obtain reliable estimates for counties with missing or small sample sizes, we followed standard procedure and combined BRFSS data from 2011 to 2019 into 3 periods of 3 years each (2011 to 2013, 2014 to 2016, 2017 to 2019). Because these factors make up 2 dimensions of the SVI, the association of SVI score with smoking prevalence by rurality and Appalachian status in our study warrants further investigation to identify the role each dimension of the SVI plays in the elevated smoking prevalence in rural Appalachia.

# Summary A (lexRankr)

Although smoking has been decreasing over the past decade, residents in rural Appalachian counties of Virginia exhibit alarmingly high rates of cigarette smoking, which likely will result in remarkable and yet preventable health and economic consequences. We analyzed the impact of SVI level and the combination of rurality with Appalachian status on county-level cigarette smoking prevalence, controlling for factors such as coal mining, tobacco agriculture, and health care provider shortages using multivariate regression analysis with robust SEs. We mapped SVI with estimated current smoking prevalence along with rural identifiers on Virginia counties using Census shape data files. Despite rural Appalachian counties having the highest cigarette smoking rates in Virginia, they did not have the highest SVI levels.

# Summary C

Paraphraser

# <https://www.paraphraser.io/text-summarizer>

# Although the prevalence of current cigarette smoking among US adults has decreased over the past several decades to 13. The Appalachian region, which extends across 13 states, has historically been characterized by its mountainous terrain, poverty, limited health care access, and reliance on tobacco agriculture and coal mining, which may contribute to elevated smoking rates. These components are closely tied to a population`s health care access and adherence to health guidelines, which may affect the effectiveness of tobacco control initiatives in rural areas with scarce resources or areas with a high SVI.

# Summary D

Ahrefs

<https://ahrefs.com/writing-tools/summarizer>

Cigarette smoking is the leading cause of preventable morbidity and mortality in the United States. Smoking rates have decreased in the US overall, but not as much in rural areas like rural Appalachia. Factors such as poverty, limited healthcare access, and reliance on tobacco agriculture and coal mining contribute to higher smoking rates in rural Appalachia.

# Summary E

Resoomer

<https://resoomer.com/en/>

Although the prevalence of current cigarette smoking among US adults has decreased over the past several decades to 13.7% in 2018, this decrease has not been as pronounced in rural areas, such as rural Appalachia . The Appalachian region, which extends across 13 states, has historically been characterized by its mountainous terrain, poverty, limited health care access, and reliance on tobacco agriculture and coal mining, which may contribute to elevated smoking rates.  Smokers living in rural Appalachia are more likely to smoke earlier in life and to smoke more heavily, and they are less likely to successfully quit.

# Summary F

TLDR This

<https://www.tldrthis.com/>

The context discusses cigarette smoking rates in Virginia counties using data from 2011-2019. It finds that rural counties, especially rural Appalachian counties, have significantly higher smoking rates than urban counties. Counties with higher social vulnerability index (SVI) scores, which measures factors like socioeconomic status, also tend to have higher smoking rates.

(will be different each timke as it is abstractive)

# Summary G

Good Content Semrush

<https://www.semrush.com/goodcontent/summary-generator/>

The prevalence of cigarette smoking among US adults has decreased over the past several decades. However, smokers in rural Appalachia are more likely to smoke earlier in life and to smoke more heavily. The social vulnerability index (SVI) evaluates 4 spheres of influence on health.

# Summary H

Text Compactor

<https://www.textcompactor.com/>

(put slider at 11%)

Specifically, we hypothesized that rural counties, Appalachian counties, and counties with higher SVI scores would have higher cigarette use prevalence. Virginia comprises 133 counties, including 80 (60.15%) urban and 53 (39.85%) rural counties, of which 25 (18.8%) are in the Appalachian region (including 17 rural Appalachian and 8 urban Appalachian). Overall cigarette smoking prevalence in Virginia was 14.80%, whereas cigarette use prevalence among rural counties was 19.38% compared with 13.20% in urban counties (difference=6.16 percentage points; Of these, the majority were classified as rural (71.43% rural and 35.80% rural Appalachian).

# Summary I

Summary Generator

<https://summarygenerator.io/>

(will vary as it is abstractive)

Cigarette smoking is a major health concern in the United States, with rural Appalachia experiencing particularly high rates. Factors such as poverty, limited healthcare access, and reliance on tobacco and coal industries contribute to elevated smoking rates in this region. The social vulnerability index (SVI) has been used to assess the impact of socioeconomic and environmental factors on health, but its association with smoking in rural areas, especially in Appalachia, is not well understood.

# Summary J (my summary)

Consequently, smokers in Appalachia are disproportionately affected by smoking-related illnesses. Specifically, we hypothesized that rural counties, Appalachian counties, and counties with higher SVI scores would have higher cigarette use prevalence. Figure 1 demonstrates the association of high SVI level with high cigarette smoking rates in counties across Virginia for the period of 2017 to 2019.

# Python rouge-score program

from rouge\_score import rouge\_scorer

scorer = rouge\_scorer.RougeScorer(['rouge1'], use\_stemmer=True)

#gold is the online sumamries in succession

#a is the lexRankr summary

#b is the rwa summary

scores = scorer.score(b, gold)

scores = scorer.score(a, gold)

scores = scorer.score(b, gold2)

scores = scorer.score(a, gold2)

scores = scorer.score(b, gold3)

scores = scorer.score(a, gold3)

scores = scorer.score(b, gold4)

scores = scorer.score(a, gold4)

scores = scorer.score(b, gold5)

scores = scorer.score(a, gold5)

Results

A

Gold1

Score(precision=0.45161290322580644, recall=0.3230769230769231,fmeasure=0.37668161434977576)

gold2

Score(precision=0.5, recall=0.2153846153846154, fmeasure=0.3010752688172043)

Gold3

Score(precision=0.5306122448979592, recall=0.4, fmeasure=0.456140350877193)

Gold4

Score(precision=0.40384615384615385, recall=0.16153846153846155,fmeasure=0.2307692307692308)

Gold5

Score(precision=0.6, recall=0.2076923076923077, fmeasure=0.30857142857142855)

Gold6

Score(precision=0.3, recall=0.23076923076923078, fmeasure=0.2608695652173913)

Gold7

Score(precision=0.4594594594594595, recall=0.26153846153846155, fmeasure=0.3333333333333333)

Gold8

precision=0.5357142857142857, recall=0.23076923076923078, fmeasure=0.32258064516129037

B

Gold1

Score(precision=0.41935483870967744, recall=0.3305084745762712, fmeasure=0.3696682464454976)

Gold2

Score(precision=0.5714285714285714, recall=0.2711864406779661, fmeasure=0.367816091954023)

Gold3

Score(precision=0.4387755102040816, recall=0.3644067796610169, fmeasure=0.39814814814814814)

Gold4

Score(precision=0.5, recall=0.22033898305084745, fmeasure=0.3058823529411765)

Gold5

Score(precision=0.5111111111111111, recall=0.19491525423728814, fmeasure=0.2822085889570552)

Gold6

Score(precision=0.39, recall=0.3305084745762712, fmeasure=0.35779816513761464)

Gold7

Score(precision=0.4864864864864865, recall=0.3050847457627119, fmeasure=0.37500000000000006)

Gold8

precision=0.5535714285714286, recall=0.2627118644067797, fmeasure=0.35632183908045983

# T test

<https://www.statology.org/minimum-sample-size-for-t-test/>

Independence: The observations in each sample should be independent.

Random Sampling: The observations in each sample should be be collected using a random sampling method.

Normality: Each sample should be roughly normally distributed.

Equal Variance: Each sample should have approximately the same variance.

# Questions

1. Is each summary satisfactory?
2. How would you rank the two?
3. Did either or both summaries miss the point of the essay?
4. Did both products present what you consider a key concept?
5. What other insights can you provide?