novis-i-hwk3-1

February 24, 2025

0.1 ECON 470 Hwk2-3

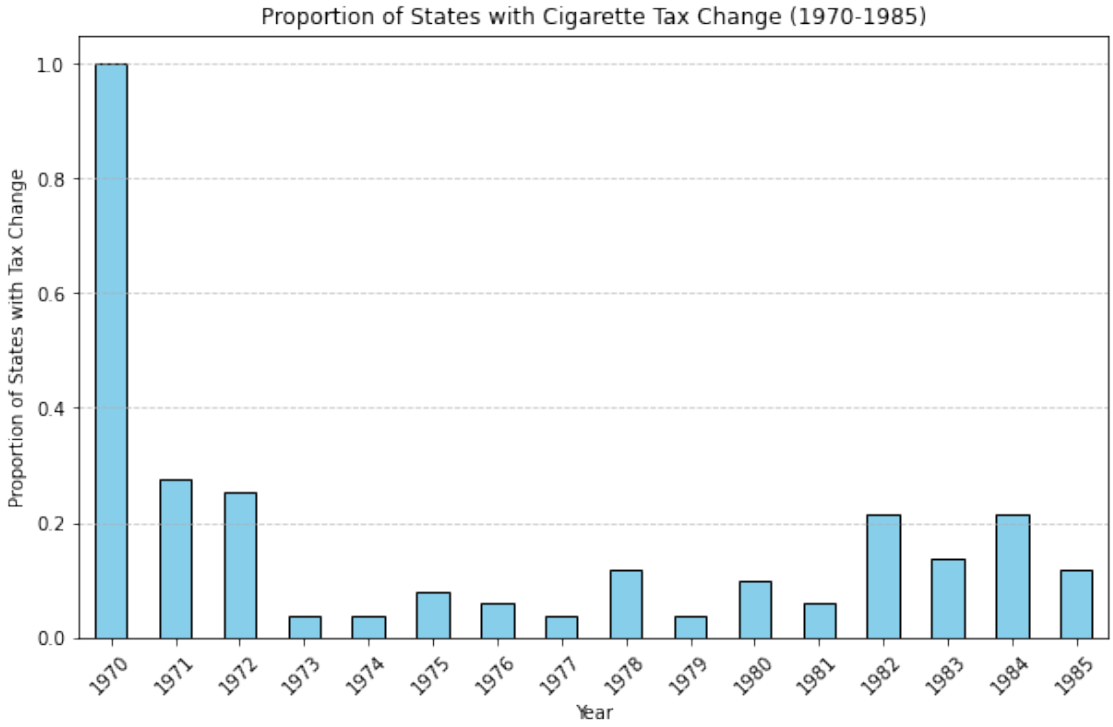
Author: Ilse Novis

Date: 3/17/2025

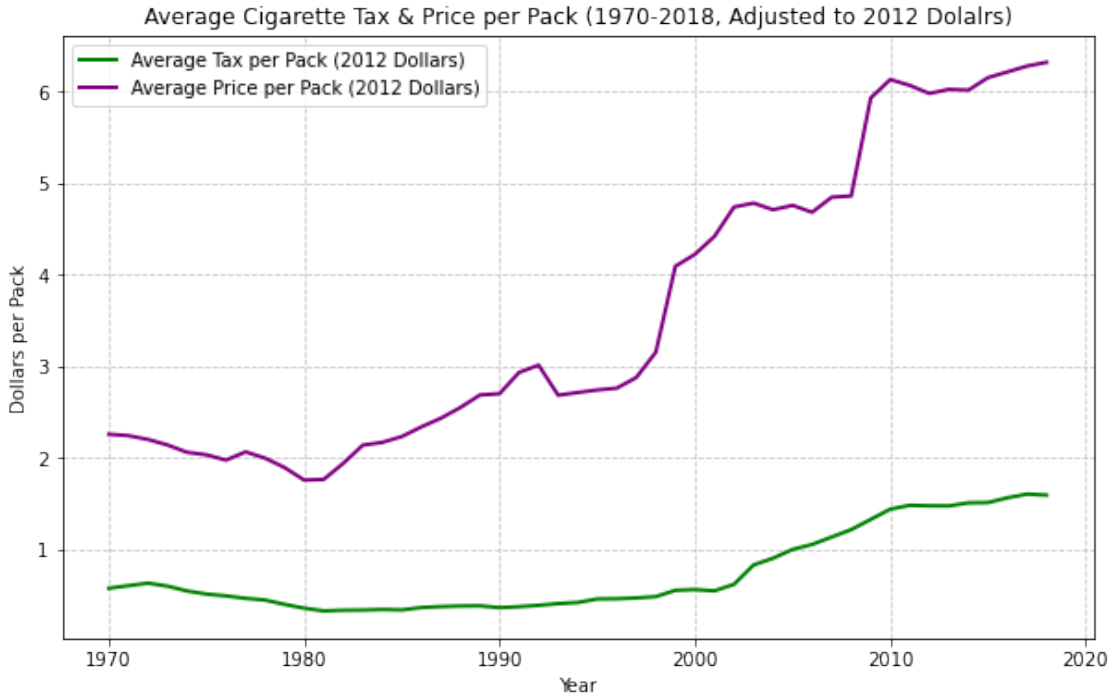
GitHub Repository

1 Summarize the Data

1.0.1 Question 1: Present a bar graph showing the proportion of states with a change in their cigarette tax in each year from 1970 to 1985.

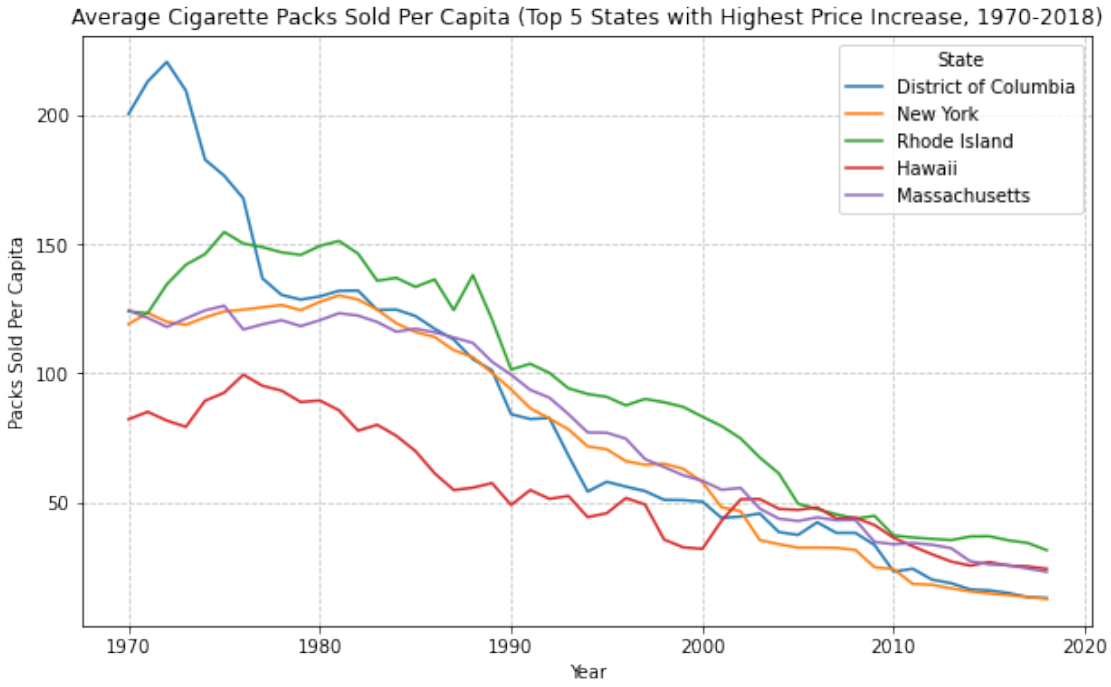


1.0.2 Question 2: Plot on a single graph the average tax (in 2012 dollars) on cigarettes and the average price of a pack of cigarettes from 1970 to 2018.



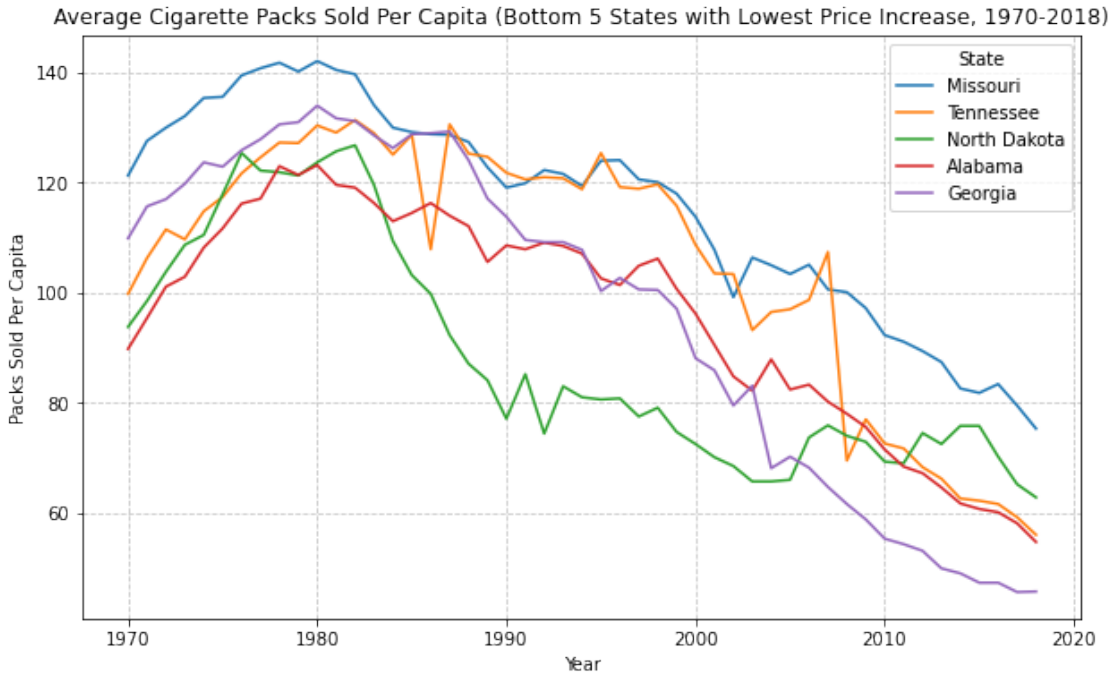
1.0.3 Question 3: Identify the 5 states with the highest increases in cigarette prices (in dollars) over the time period. Plot the average number of packs sold per capita for those states from 1970 to 2018.

```
Top 5 states with highest pring increase: ['District of Columbia', 'New York',
'Rhode Island', 'Hawaii', 'Massachusetts']
```



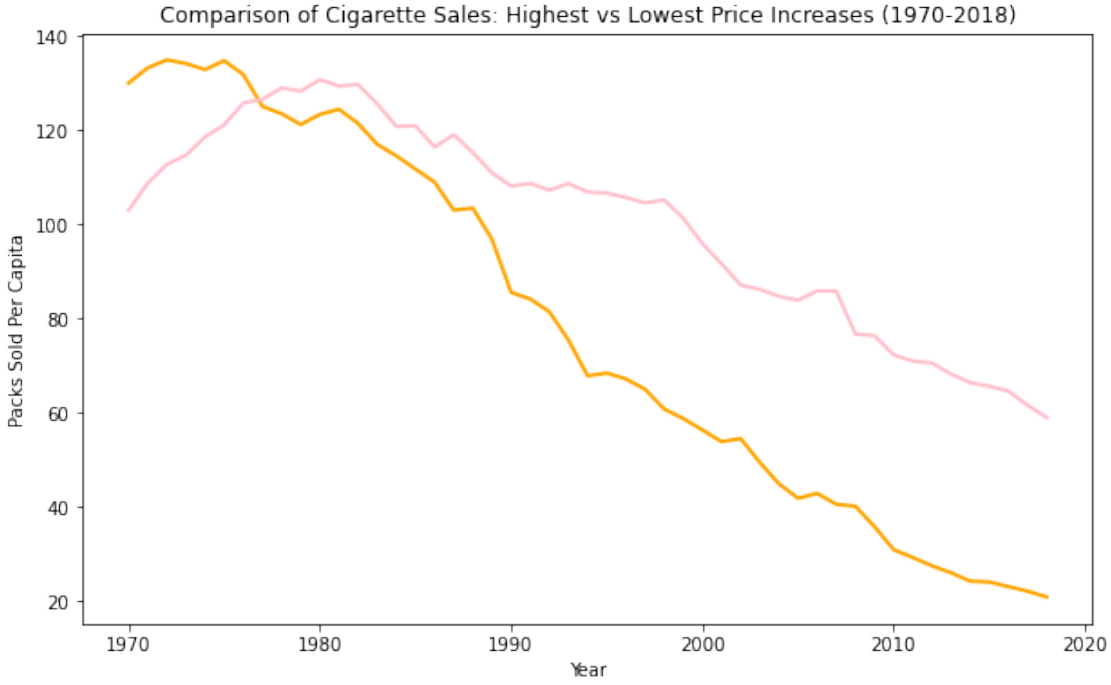
1.0.4 Question 4: Identify the 5 states with the lowest increases in cigarette prices over the time period. Plot the average number of packs sold per capita for those states from 1970 to 2018.

```
Bottom 5 states with lowest price increase: ['Missouri', 'Tennessee', 'North
Dakota', 'Alabama', 'Georgia']
```



1.0.5 Question 5: Compare the trends in sales from the 5 states with the highest price increases to those with the lowest price increases.

```
Top 5 states with highest price increase: ['District of Columbia', 'New York',
'Rhode Island', 'Hawaii', 'Massachusetts']
Bottom 5 states with lowest price increase: ['Missouri', 'Tennessee', 'North
Dakota', 'Alabama', 'Georgia']
```



2 Estimate ATEs

2.0.1 Question 6: Focusing only on the time period from 1970 to 1990, regress log sales on log prices to estimate the price elasticity of demand over that period. Interpret your results.

OLS Regression Results

Dep. Variable:	log_sales	R-squared:	0.294
Model:	OLS	Adj. R-squared:	0.293
Method:	Least Squares	F-statistic:	445.1
Date:	Mon, 24 Feb 2025	Prob (F-statistic):	6.98e-83
Time:	17:00:39	Log-Likelihood:	263.40
No. Observations:	1071	AIC:	-522.8
Df Residuals:	1069	BIC:	-512.8
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	5.4288	0.030	182.065	0.000	5.370	5.487
log_price	-0.8094	0.038	-21.098	0.000	-0.885	-0.734

Omnibus:	89.160	Durbin-Watson:	0.183
Prob(Omnibus):	0.000	Jarque-Bera (JB):	466.536
Skew:	0.128	Prob(JB):	4.93e-102
Kurtosis:	6.223	Cond. No.	10.5

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Estimated Price Elasticity of Demand: -0.809
The estimated price elasticity of demand is -0.809, meaning that a 1% increase in price is associated with a 0.8% decrease in cigarette consumption.

2.0.2 Question 7: Regress log sales on log prices using the total (federal and state) cigarette tax (in dollars) as an instrument for log prices. Interpret your results and compare your estimates to those without an instrument. Are they different? If so, why?

Question 7.1:

First-Stage Regression Results:

OLS Regression Results

Dep. Variable:	log_price	R-squared:	0.617
Model:	OLS	Adj. R-squared:	0.617
Method:	Least Squares	F-statistic:	1725.
Date:	Mon, 24 Feb 2025	Prob (F-statistic):	2.80e-225
Time:	17:00:39	Log-Likelihood:	1020.7
No. Observations:	1071	AIC:	-2037.
Df Residuals:	1069	BIC:	-2027.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.2355	0.012	105.227	0.000	1.212	1.259
log_tax	0.3328	0.008	41.537	0.000	0.317	0.349

Omnibus:	6.850	Durbin-Watson:	0.303
Prob(Omnibus):	0.033	Jarque-Bera (JB):	5.505
Skew:	0.081	Prob(JB):	0.0638
Kurtosis:	2.689	Cond. No.	8.72

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Question 7.2:

Second-Stage (IV) Regression Results:

OLS Regression Results

Dep. Variable:	log_sales	R-squared:	0.236
Model:	OLS	Adj. R-squared:	0.235
Method:	Least Squares	F-statistic:	330.3
Date:	Mon, 24 Feb 2025	Prob (F-statistic):	1.56e-64
Time:	17:00:39	Log-Likelihood:	221.17
No. Observations:	1071	AIC:	-438.3
Df Residuals:	1069	BIC:	-428.4
Df Model:	1		

```
Covariance Type:            nonrobust
================================================================================
=
                    coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------------
-
const             5.5155      0.039    140.744      0.000       5.439
5.592
log_price_hat    -0.9231      0.051    -18.175      0.000      -1.023
-0.823
==============================================================================
Omnibus:                       83.338   Durbin-Watson:                   0.157
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              430.014
Skew:                           0.023   Prob(JB):                     4.20e-94
Kurtosis:                       6.104   Cond. No.                         13.4
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.

IV-Estimated Price Elasticity of Demand: -0.923

OLS-Estimated Price Elasticity: -0.809
Difference between OLS and IV Estimates: 0.114
```

Question 7: Interpretation

The IV estimate is more negative (elastic), suggesting OLS underestimated price elasticity due to endogeneity.

2.0.3 Question 8: Show the first stage and reduced-form results from the instrument.

Question 8.1: First Stage Form

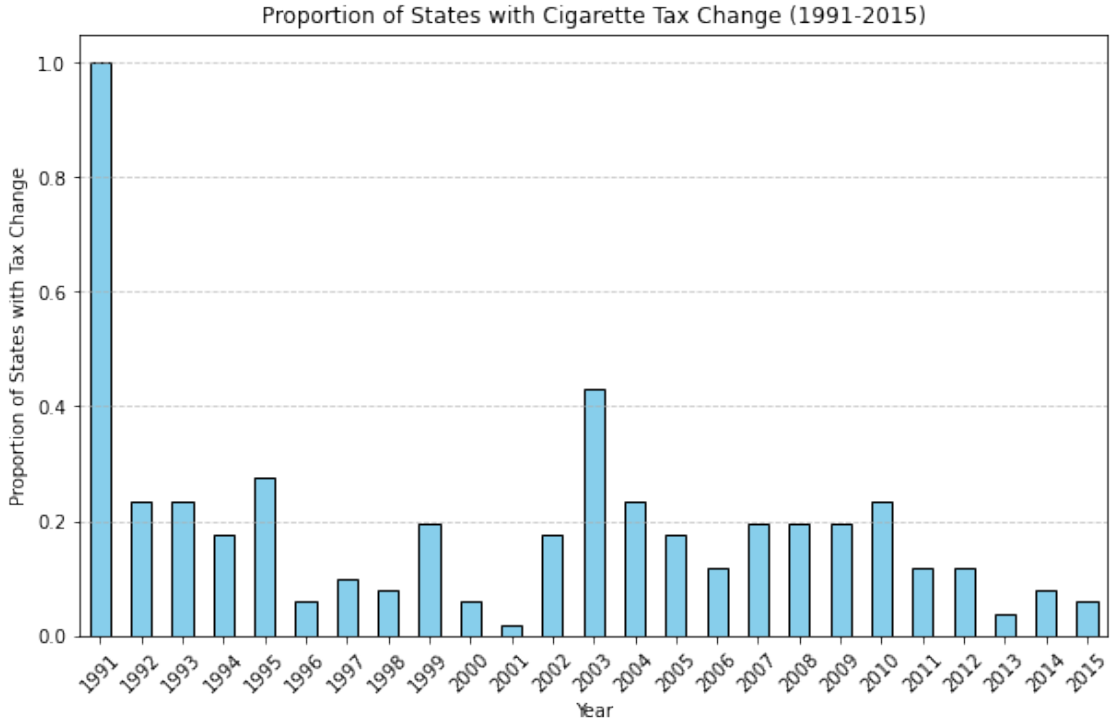
```
=== First-Stage Regression: Log(Price) ~ Log(Tax) ===
                            OLS Regression Results
==============================================================================
Dep. Variable:              log_price   R-squared:                       0.617
Model:                            OLS   Adj. R-squared:                  0.617
Method:                 Least Squares   F-statistic:                     1725.
Date:                Mon, 24 Feb 2025   Prob (F-statistic):          2.80e-225
Time:                        17:00:39   Log-Likelihood:                 1020.7
No. Observations:                1071   AIC:                            -2037.
Df Residuals:                    1069   BIC:                            -2027.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          1.2355      0.012    105.227      0.000       1.212       1.259
log_tax        0.3328      0.008     41.537      0.000       0.317       0.349
==============================================================================
Omnibus:                        6.850   Durbin-Watson:                   0.303
Prob(Omnibus):                  0.033   Jarque-Bera (JB):                5.505
Skew:                           0.081   Prob(JB):                       0.0638
Kurtosis:                       2.689   Cond. No.                         8.72
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

Question 8.2: Reduced Form Regression

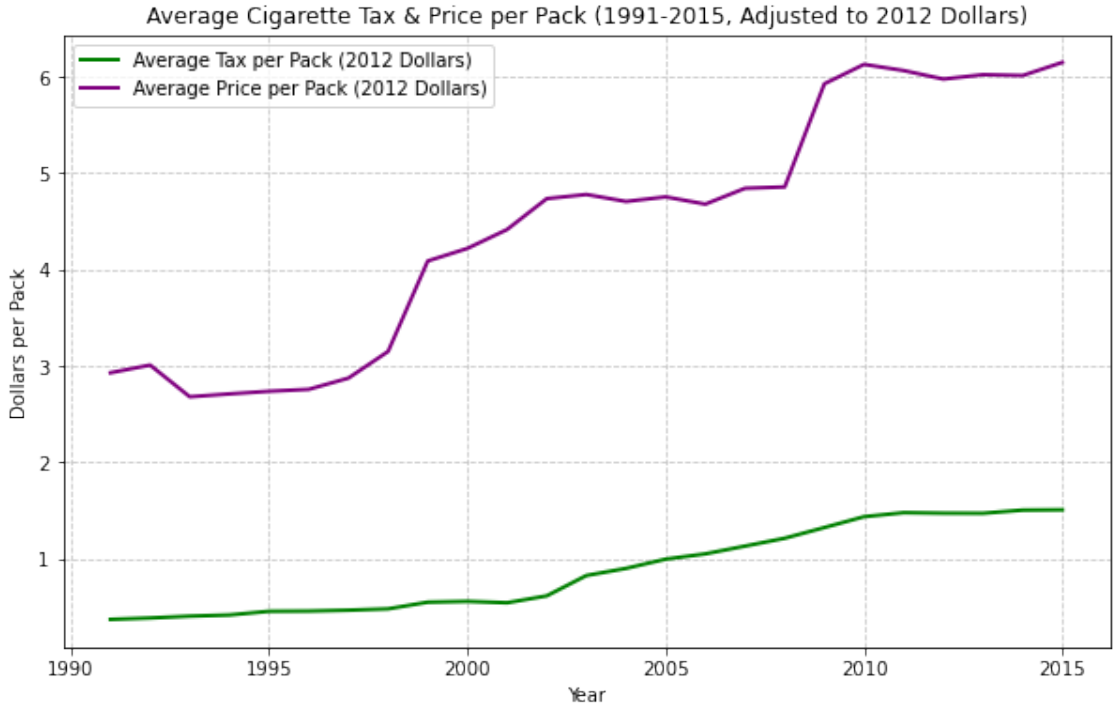
```
=== Reduced-Form Regression: Log(Sales) ~ Log(Tax) ===
                            OLS Regression Results
==============================================================================
Dep. Variable:              log_sales   R-squared:                       0.236
Model:                            OLS   Adj. R-squared:                  0.235
Method:                 Least Squares   F-statistic:                     330.3
Date:                Mon, 24 Feb 2025   Prob (F-statistic):           1.56e-64
Time:                        17:00:39   Log-Likelihood:                 221.17
No. Observations:                1071   AIC:                            -438.3
Df Residuals:                    1069   BIC:                            -428.4
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          4.3750      0.025    176.627      0.000       4.326       4.424
log_tax       -0.3072      0.017    -18.175      0.000      -0.340      -0.274
==============================================================================
Omnibus:                       83.338   Durbin-Watson:                   0.157
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              430.014
Skew:                           0.023   Prob(JB):                     4.20e-94
Kurtosis:                       6.104   Cond. No.                         8.72
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
```

2.0.4 Question 9: Repeat questions 1-3 focusing on the period from 1991 to 2015.

Question 9.1: Bar graph showing the proportion of states with a change in their cigarette tax

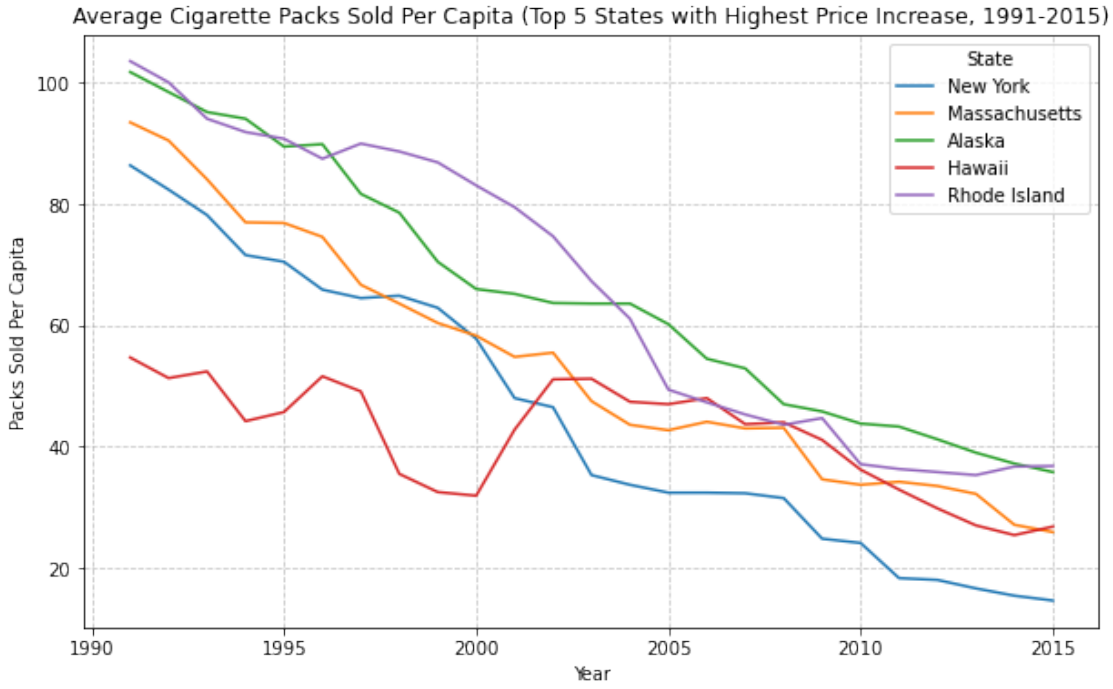


Question 9.2: Average tax (in 2012 dollars) on cigarettes and the average price of a pack of cigarettes from 1991 to 2015



Question 9.3: Identify the 5 states with the highest increases in cigarette prices (in dollars) over the time period. Plot the average number of packs sold per capita for those states from 1991 to 2015

```
Top 5 states with highest price increase (1991-2015): ['New York',
'Massachusetts', 'Alaska', 'Hawaii', 'Rhode Island']
```



2.0.5 Question 10: Compare your elasticity estimates from 1970-1990 versus those from 1991-2015. Are they different? If so, why?

```
=== Price Elasticity Estimate for 1970-1990 ===
                            OLS Regression Results
==============================================================================
Dep. Variable:              log_sales   R-squared:                       0.294
Model:                            OLS   Adj. R-squared:                  0.293
Method:                 Least Squares   F-statistic:                     445.1
Date:                Mon, 24 Feb 2025   Prob (F-statistic):           6.98e-83
Time:                        17:00:40   Log-Likelihood:                 263.40
No. Observations:                1071   AIC:                            -522.8
Df Residuals:                    1069   BIC:                            -512.8
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          5.4288      0.030    182.065      0.000       5.370       5.487
log_price     -0.8094      0.038    -21.098      0.000      -0.885      -0.734
==============================================================================
Omnibus:                       89.160   Durbin-Watson:                   0.183
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              466.536
Skew:                           0.128   Prob(JB):                    4.93e-102
Kurtosis:                       6.223   Cond. No.                         10.5
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Estimated Price Elasticity: -0.809

=== Price Elasticity Estimate for 1991-2015 ===
                            OLS Regression Results
==============================================================================
Dep. Variable:              log_sales   R-squared:                       0.561
Model:                            OLS   Adj. R-squared:                  0.561
Method:                 Least Squares   F-statistic:                     1630.
Date:                Mon, 24 Feb 2025   Prob (F-statistic):          4.20e-230
Time:                        17:00:40   Log-Likelihood:                -256.00
No. Observations:                1275   AIC:                             516.0
Df Residuals:                    1273   BIC:                             526.3
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const          5.6617      0.036    155.427      0.000       5.590       5.733
log_price     -0.9968      0.025    -40.370      0.000      -1.045      -0.948
==============================================================================
Omnibus:                       23.003   Durbin-Watson:                   0.208
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               43.688
Skew:                           0.011   Prob(JB):                     3.26e-10
Kurtosis:                       3.907   Cond. No.                         9.35
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
Estimated Price Elasticity: -0.997

=== Elasticity Comparison ===
Elasticity (1970-1990): -0.809
Elasticity (1991-2015): -0.997
Difference: 0.187


Demand became more elastic (greater sensitivity to price changes) in 1991-2015.
```