Write clear and complete answers. Partial answers get partial credit. Show your work to get credit. Be sure to describe the steps taken to reach your conclusions.

1. Take a regression of votes by individual precincts (voting units) in Palm Beach, Florida.

\[
\text{Overvotes} = -3.44 - 0.017 \text{Age}_{18-29} + 0.041 \text{Age}_{65} + 0.089 \text{Black} \\
\begin{array}{ll}
(3.06) & (0.028) & (0.007) & (0.010) \\
\end{array}
\]

\[
+ 0.205 \text{Hispanic} + 0.064 \text{Female} + 0.099 \text{Democrat} \\
\begin{array}{llll}
(0.038) & (0.046) & (0.034) \\
\end{array}
\]

\[
+ 0.028 \text{OtherParty} - 0.044 \text{Nelson} \\
\begin{array}{ll}
(0.064) & (0.032) \\
\end{array}
\]

\[
R^2 = 0.71 \quad \hat{\sigma} = 1.81 \quad SSR = 1595 \quad n = 494
\]

The data is DataSet #1, described at the end of the exam. Test the hypothesis that \( \text{Age}_{65} \) has no effect on Overvotes, against the alternative that it has an effect. Use the 95% significance level. Does this result make sense?

2. Take a regression of votes by individual precincts (voting units) in Palm Beach, Florida.

\[
\text{Gore - Bush} = -110.1 - 0.071 \text{Age}_{18-29} - 0.040 \text{Age}_{65} - 0.035 \text{Black} \\
\begin{array}{llll}
(3.9) & (0.084) & (0.022) & (0.023) \\
\end{array}
\]

\[
- 0.182 \text{Hispanic} + 0.045 \text{Female} + 0.306 \text{Democrat} \\
\begin{array}{llll}
(0.107) & (0.081) & (0.084) \\
\end{array}
\]

\[
+ 0.632 \text{OtherParty} + 1.771 \text{Nelson} - 0.578 \text{Overvotes} \\
\begin{array}{llll}
(0.116) & (0.076) & (0.137) \\
\end{array}
\]

\[
R^2 = 0.98 \quad \hat{\sigma} = 4.09 \quad SSR = 8086 \quad n = 494
\]

The data is DataSet #1, described at the end of the exam. Form a 95% confidence interval for the coefficient of Overvotes in this regression. Assume that this coefficient can be interpreted as the effect of an “overvote” on the margin between Gore and Bush. What do you conclude?
3. Take a regression of votes by individual precincts (voting units) in Palm Beach, Florida. You estimate two regressions for the percentage of votes received by Bush. The first is

\[
Bush = 104.4 + 0.049 \text{ Age}_{18-29} + 0.015 \text{ Age}_{65} + 0.004 \text{ Black} \\
(2.2) \quad (0.041) \quad (0.011) \quad (0.009)
\]

\[
+ 0.039 \text{ Hispanic} - 0.040 \text{ Female} - 0.174 \text{ Democrat} \\
(0.045) \quad (0.040) \quad (0.040)
\]

\[- 0.321 \text{ OtherParty} - 0.876 \text{ Nelson} \\
(0.059) \quad (0.036)
\]

\[
R^2 = 0.983018 \quad \hat{\sigma} = 2.0631 \quad SSR = 2064 \quad n = 494
\]

The second is

\[
Bush = -102.9 - 0.158 \text{ Democrat} - 0.299 \text{ OtherParty} - 0.884 \text{ Nelson} \\
(0.8) \quad (0.029) \quad (0.039) \quad (0.027)
\]

\[
R^2 = 0.982732 \quad \hat{\sigma} = 2.0697 \quad SSR = 2099 \quad n = 494
\]

The data is DataSet #1, described at the end of the exam. Use this information to test the hypothesis that the demographic variables \text{Age}_{18-29}, \text{Age}_{65}, \text{Black}, \text{Hispanic}, \text{Female} do not enter the equation for Bush’s vote. Use a 5\% significance value.

4. Take a regression of votes by county in Florida, for all counties excluding Palm Beach.

\[
Buchanan = 0.0123 - 0.019 (\text{Age}_{65} - 23.7) - 0.006 (\text{Black} - 14.4) \\
(0.0557) \quad (0.004) \quad (0.003)
\]

\[- 0.012 (\text{Hispanic} - 9.8) - 0.011 (\text{College} - 22.1) \\
(0.004) \quad (0.004)
\]

\[- 0.026 (\text{Income} - 33.5) \\
(0.007)
\]

\[
R^2 = 0.49 \quad \hat{\sigma} = 0.24 \quad SSR = 3.46 \quad n = 66
\]

The data is DataSet #2, described at the end of the exam. From each regressor we have subtracted the value of that regressor in Palm Beach. For example, 23.7 is the percentage of the population in Palm Beach over the age of 65.

Using the information in this regression, perform an out-of-sample forecast of the percentage of votes received by Buchanan in Palm Beach. Calculate the point forecast. Calculate a 95\% forecast confidence interval.

The actual percentage Buchanan received in Palm Beach was 0.79\%. Feel free to (briefly) comment.
5. Assume that you have data \( \{y_{1i}, y_{2i}, x_i\} \) for random individuals \( i = 1, \ldots, n \). Assume that the variables satisfy the two linear regressions

\[
\begin{align*}
y_{1i} &= \alpha_0 + \alpha_1 x_i + u_{1i} \\
y_{2i} &= \beta_0 + \beta_1 x_i + u_{2i}
\end{align*}
\]

where

\[
\begin{align*}
E (u_{1i} \mid x_i) &= 0 \\
E (u_{2i} \mid x_i) &= 0.
\end{align*}
\]

How can you test the hypothesis that \( \alpha_1 = \beta_1 \)? Describe the regression(s) you would estimate, what would be the test statistic, and what would lead to reject of the null hypothesis.
DataSet #1. Votes by Precinct (voting unit) in Palm Beach, Florida.

- Gore: Percentage of Votes cast for President marked for “Al Gore”
- Bush: Percentage of Votes cast for President marked for “George Bush”
- Gore-Bush: Difference between above
- Overvotes: Percentage of Votes cast for President marked with two or more candidates
- Age18-29: Percentage of Registered voters (in precinct) between ages of 18 and 29
- Age65: Percentage of Registered voters over age of 65
- Black: Percentage of Registered voters
- Hispanic: Percentage of Registered voters, any race
- Female: Percentage of Registered voters
- Democrat: Percentage of Registered voters who belong to Democrat Party
- OtherParty: Percentage of Registered voters who belong to a political party other than Democrat or Republican
- Nelson: Percentage of Votes cast for Senator marked for “Bill Nelson”, the Democrat candidate for Senate

DataSet #2. Votes by County in State of Florida

- Buchanan: Percentage of Votes cast for President marked for “Pat Buchanan”
- Age65: Percentage of population in county over age of 65 (U.S. Census)
- Black: Percentage of population in county (U.S. Census)
- Hispanic: Percentage of population in county (U.S. Census)
- College: Percentage of population in county who are college-educated (U.S. Census)
- Income: Median Household Income in county (in thousands of dollars) (U.S. Census)