

Exercise 2: Policy Memorandum

To: Mayor of Smallville

Regarding: Smallville Sanitation Management

From: Yehzee Ryoo

Executive summary

Resolving the current mismanagement of the Sanitation Department in Smallville should be one of the first priorities as a new elected mayor to increase the quality of life of your constituencies. You can consider eliminating the Sanitation department and switching to a private contractor which offers a much lower price – \$40.60 – compared to the current cost – \$48.50 – the department has used per household for once-a-week curbside pick-up. However, this solution can be politically risky since many of your constituencies may lose their jobs. Thus, I would like to recommend another way of addressing the current problem – appointing a more competent Sanitation Department supervisor. By adopting this solution, the cost of refuse collection can be reduced substantively while preventing any political cost that might follow a massive unemployment of the residents caused by removing the department.

Comparison with other municipalities

First of all, I would like to compare the refuse collection cost of Smallville with other 30 municipalities to evaluate if the current cost is a suitable amount. Looking at the summary table in the Appendix, the average cost per household for weekly refuse pickup of 30 municipalities in the region was around \$44.97 last year. It is clear that the Sanitation Department of Smallville is expending about \$3.53 more per household, which in total \$554,210 every week. Given that Smallville is spending more than other municipalities in general, whether this excess of cost is reasonable or not considering the various factors that determine the refuse collection cost

should be examined. Probable factors that affect the cost could include the number of households in the city, the density of households, the average wage of collection workers, and the number of snow emergencies that can increase the costs of collection. Incorporating all these factors, I develop a statistical model and predict the appropriate amount that your city should be spending for the refuse collection.

According to my model, the appropriate amount of the collection cost for a city with factor values that your city has seems to be \$37.524923. This was calculated by the multivariate regression model having four aforementioned factors – the number of households, the density of households, the average wage of collection workers, and the number of snow emergencies – and a square value of one of them, the number of households, as independent variables. The squared variable was included after I examined a plot of residuals, since it was showing a certain pattern when it should be evenly distributed. Following this result, Smallville can save around \$10.98 with a better management of the Sanitation Department.

Recommendation of reappointment

To conclude, my model suggests that Smallville can reduce the cost of collection with a better management without eliminating the current department. The expected cost with a new supervisor is around \$37.52, which is even lower than the potential bid from the private contractor. Even though there is a range of variation in this computed cost, political backlash from your residents could be very risky since you were elected very recently. Thus, I recommend you promoting a better management of the sanitation services by hiring a new, more competent supervisor to the Sanitation department, rather than switching to a private contractor.

Appendix

```
. summarize cost_per_household hholds density wage snowdays
```

Variable	Obs	Mean	Std. Dev.	Min	Max
cost_per_h~d	30	44.96533	16.27534	20.69	84.15
hholds	30	7.563067	2.843379	2.56	12.23
density	30	607.6667	83.15088	458.8	743.8
wage	30	17.87533	1.90663	15.61	21.36
snowdays	30	4.466667	2.849481	0	10

```
. regress cost_per_household hholds density wage snowdays
```

Source	SS	df	MS	Number of obs	=	30
				F(4, 25)	=	17.47
Model	5657.89948	4	1414.47487	Prob > F	=	0.0000
Residual	2023.81858	25	80.9527431	R-squared	=	0.7365
				Adj R-squared	=	0.6944
Total	7681.71806	29	264.88683	Root MSE	=	8.9974

cost_per_h~d	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
hholds	3.63205	.7313792	4.97	0.000	2.125747 5.138354
density	-.0268314	.0211887	-1.27	0.217	-.0704702 .0168074
wage	.8352337	.9116187	0.92	0.368	-1.04228 2.712748
snowdays	1.837679	.6995717	2.63	0.015	.3968837 3.278473
_cons	10.66206	20.95837	0.51	0.615	-32.50251 53.82664

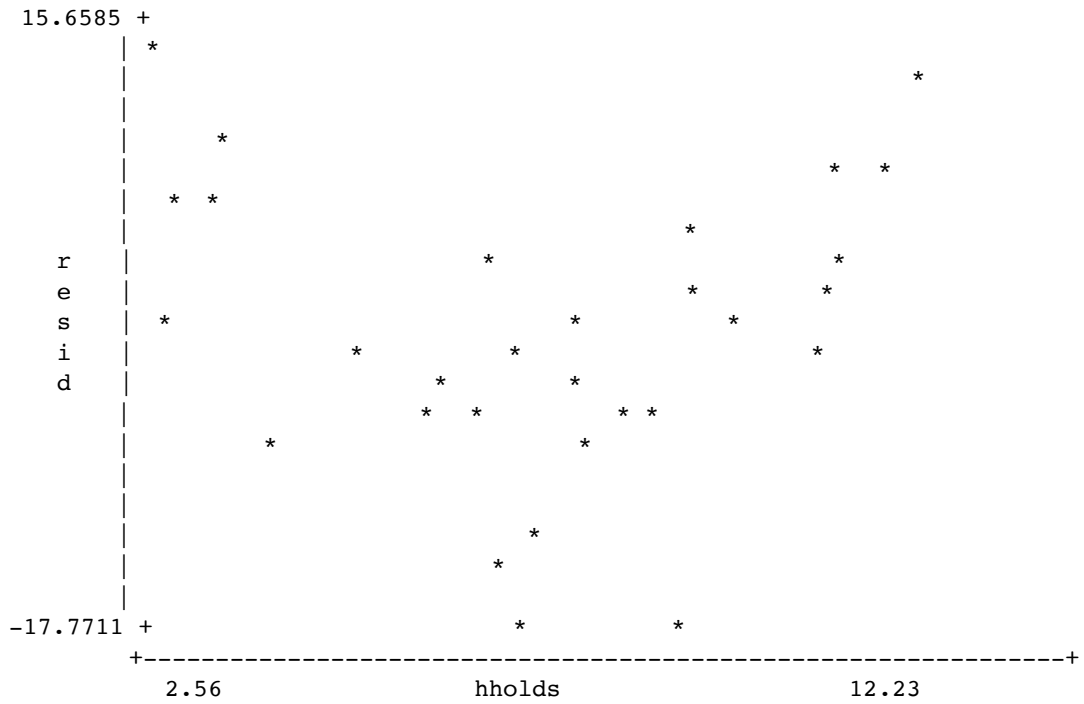
```
.
```

```

. **Plot residuals
. predict p_cost_per_household
(option xb assumed; fitted values)

.
. generate resid = cost_per_household - p_cost_per_household
.
. plot resid hholds

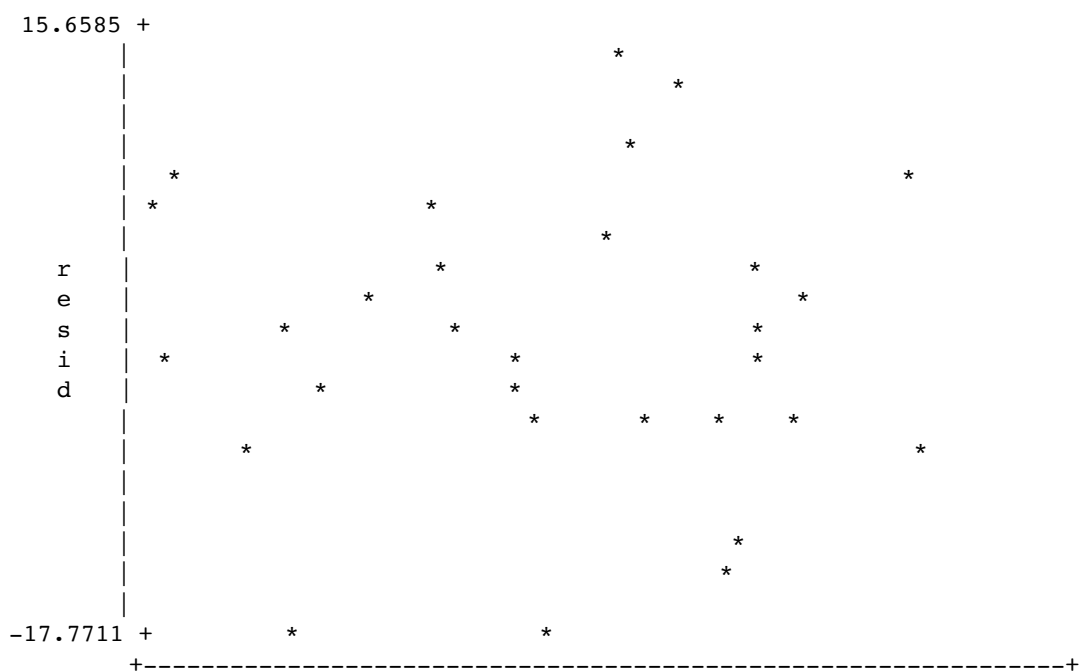
```



```

. **maybe there is a quadratic pattern?
.
. plot resid density

```

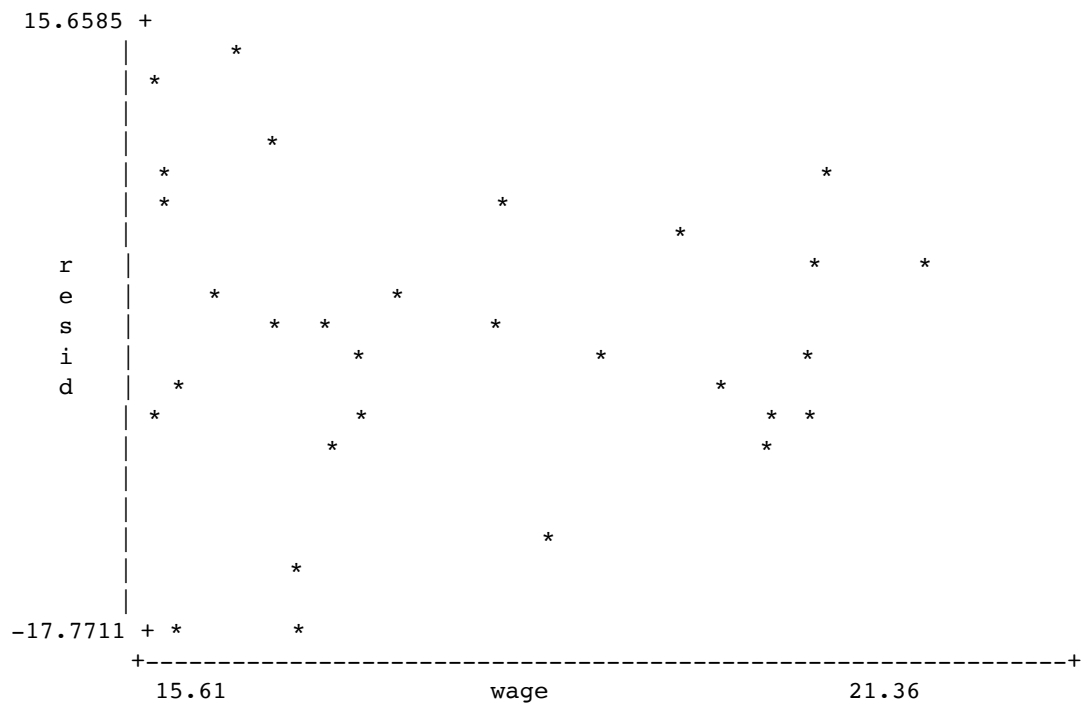


458.8

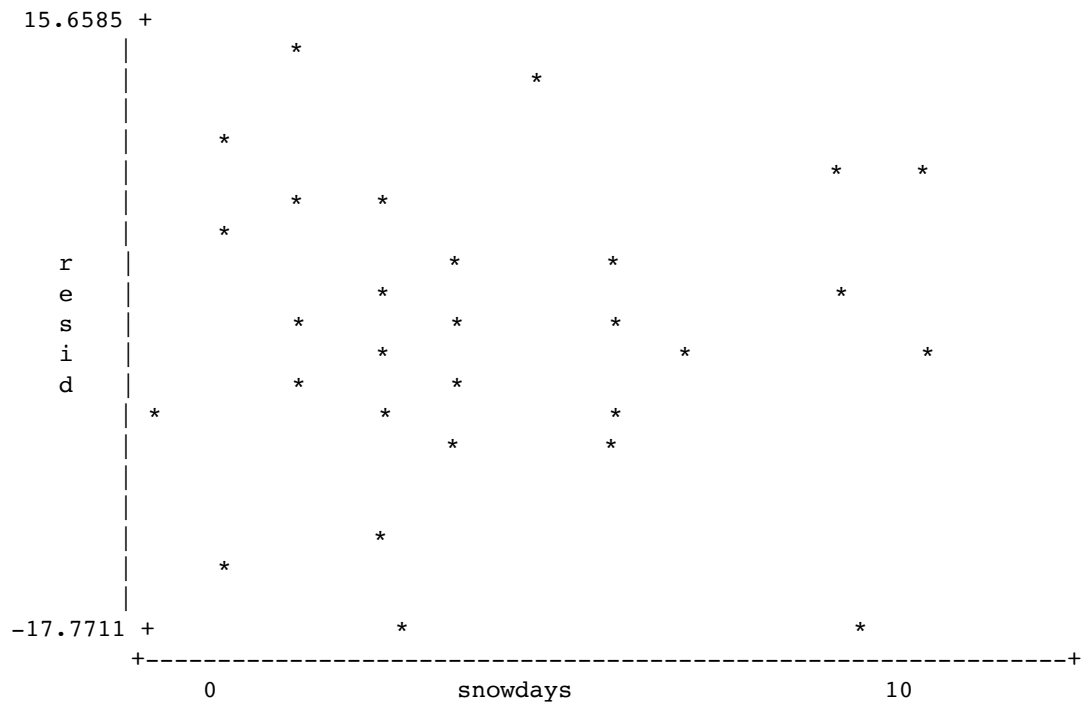
density

743.8

. plot resid wage



. plot resid snowdays



.
 . **Include squared term of hholds
 . gen hholds2 = hholds^2

```

.
. regress cost_per_household hholds hholds2 density wage snowdays

-----+-----
Source |      SS      df      MS      Number of obs      =      30
-----+-----+-----+----- F(5, 24)      =      52.29
Model | 7035.86261      5 1407.17252      Prob > F      =      0.0000
Residual | 645.855448     24 26.9106437      R-squared      =      0.9159
-----+-----+-----+----- Adj R-squared =      0.8984
Total | 7681.71806     29 264.88683      Root MSE      =      5.1875

-----+-----
cost_per_h~d |      Coef.      Std. Err.      t      P>|t|      [95% Conf. Interval]
-----+-----+-----+-----+-----+-----
hholds | -9.821759      1.926841     -5.10      0.000     -13.79856     -5.844954
hholds2 |  .9315048      .1301752      7.16      0.000      .6628364      1.200173
density | -.0038076      .0126332     -0.30      0.766     -.0298812      .022266
wage | 2.051062      .5523854      3.71      0.001      .9109951      3.19113
snowdays | .9323627      .4227229      2.21      0.037      .0599055      1.80482
_cons | 20.1717      12.15666      1.66      0.110     -4.918425     45.26182

-----+-----

.
. **Variance-Covariance model for b
. * s^2*(X'X)^-1
. * but can easily calculate with..
. matrix V = get(VCE)

. matrix list V, nohalf

symmetric V[6,6]
      hholds      hholds2      density      wage      snowdays      _cons
hholds  3.7127181  -.24474672  -.00752571  -.3536287  .15206464  -1.9521241
hholds2 -.24474672  .01694559  .00041884  .02211791  -.01646917  .17299576
density -.00752571  .00041884  .0001596   .00144048  -.00018558  -.09221544
wage    -.3536287  .02211791  .00144048  .30512959  -.04460116  -4.8938879
snowdays .15206464  -.01646917  -.00018558  -.04460116  .17869467  .03253608
_cons   -1.9521241  .17299576  -.09221544  -4.8938879  .03253608  147.7845

.
. matrix x0 = 6.28 \ 6.28^2 \ 620 \ 19.50 \ 5 \ 1

.
. matrix x0tVx0 = (x0')*V*x0

. matrix list x0tVx0

symmetric x0tVx0[1,1]
      c1
c1 2.7406849

.
. **Estimator of the variance of e (s^2)
. * we can calculate this using RSS
. * s^2 = RSS / df = 645.855448 / 24
. di 645.855448 / 24
26.910644

. * = 26.910644

```

```

.
.
. **Variance of y_hat
. * var(y0_hat) = s^2 + x0tVx0
. di 26.910644 + 2.7406849
29.651329

. * = 29.651329
.
.
. **Y0_hat
. * y0_hat = -9.821759 * 6.28 + -.0038076 * 620 + 2.051062 * 19.50 + .9323627
* 5
> + .9315048 * 6.28^2
. di -9.821759 * 6.28 + -.0038076 * 620 + 2.051062 * 19.50 + .9323627 * 5 +
.931504
> 8 * 6.28^2 + 20.1717
37.524923

. * = 37.524923
.
. **Confidence interval
. * y0_hat +/- tvalue*var(y0_hat)^1/2
. di 37.524923 - 2.064 * sqrt(29.651329)
26.285817

. di 37.524923 + 2.064 * sqrt(29.651329)
48.764029

. * from 26.285817 to 48.764029
.

```