Booklet of Figures
for
STAC32 Final Exam

You may take away this booklet after the exam, and therefore you are free to tear off pages such as the coloured graph at the end.

List of Figures in this document by page:

List of Figures

```
library(tidyverse)

## -- Attaching packages ---------------------------------
tidyverse 1.2.1 --
## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
## -- Conflicts ------------------------------------
tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(broom)
```

Figure 1: Packages

```
length ratio satisfaction
06-hours 2-1 25
06-hours 2-1 26
06-hours 2-1 28
06-hours 2-1 27
06-hours 4-1 31
06-hours 4-1 26
06-hours 4-1 29
06-hours 4-1 27
06-hours 8-1 24
06-hours 8-1 25
06-hours 8-1 28
06-hours 8-1 26
09-hours 2-1 26
09-hours 2-1 29
09-hours 2-1 27
09-hours 2-1 30
09-hours 4-1 25
09-hours 4-1 30
09-hours 4-1 24
09-hours 4-1 26
09-hours 8-1 33
09-hours 8-1 25
09-hours 8-1 28
09-hours 8-1 27
12-hours 2-1 22
12-hours 2-1 25
12-hours 2-1 20
12-hours 2-1 21
12-hours 4-1 33
12-hours 4-1 25
12-hours 4-1 27
12-hours 4-1 27
12-hours 8-1 30
12-hours 8-1 26
12-hours 8-1 31
12-hours 8-1 27
```

Figure 2: Time-of-day electricity pricing data

Obs	eating	fne
1	bulimic	21
2	bulimic	13
3	bulimic	10
4	bulimic	20
5	bulimic	25
6	bulimic	19
7	bulimic	16
8	bulimic	21
9	bulimic	24
10	bulimic	13
11	bulimic	14
12	normal	13
13	normal	6
14	normal	16
15	normal	13
16	normal	8
17	normal	19
18	normal	23
19	normal	18
20	normal	11
21	normal	19
22	normal	7
23	normal	10
24	normal	15
25	normal	20

Figure 3: FNE data

eating	N	Mean	Std Dev	Std Err	Minimum	Maximum
bulimic	11	17.8182	4.9157	1.4821	10.0000	25.0000
normal	14	14.1429	5.2894	1.4137	6.0000	23.0000
Diff (1-2)		3.6753	5.1303	2.0670		

eating	Method	Mean	95% CL Mean		Std Dev
bulimic		17.8182	14.5158	21.1206	4.9157
normal		14.1429	11.0888	17.1969	5.2894
Diff (1-2)	Pooled	3.6753	0.1327	Infty	5.1303
Diff (1-2)	Satterthwaite	3.6753	0.1602	Infty	

eating	Method	95% CL Std Dev	
bulimic		3.4346	8.6266
normal		3.8346	8.5215
Diff (1-2)	Pooled	3.9873	7.1965
Diff (1-2)	Satterthwaite		

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	23	1.78	0.0443
Satterthwaite	Unequal	22.284	1.79	0.0432

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	13	10	1.16	0.8305

Figure 4: FNE t-test, text output

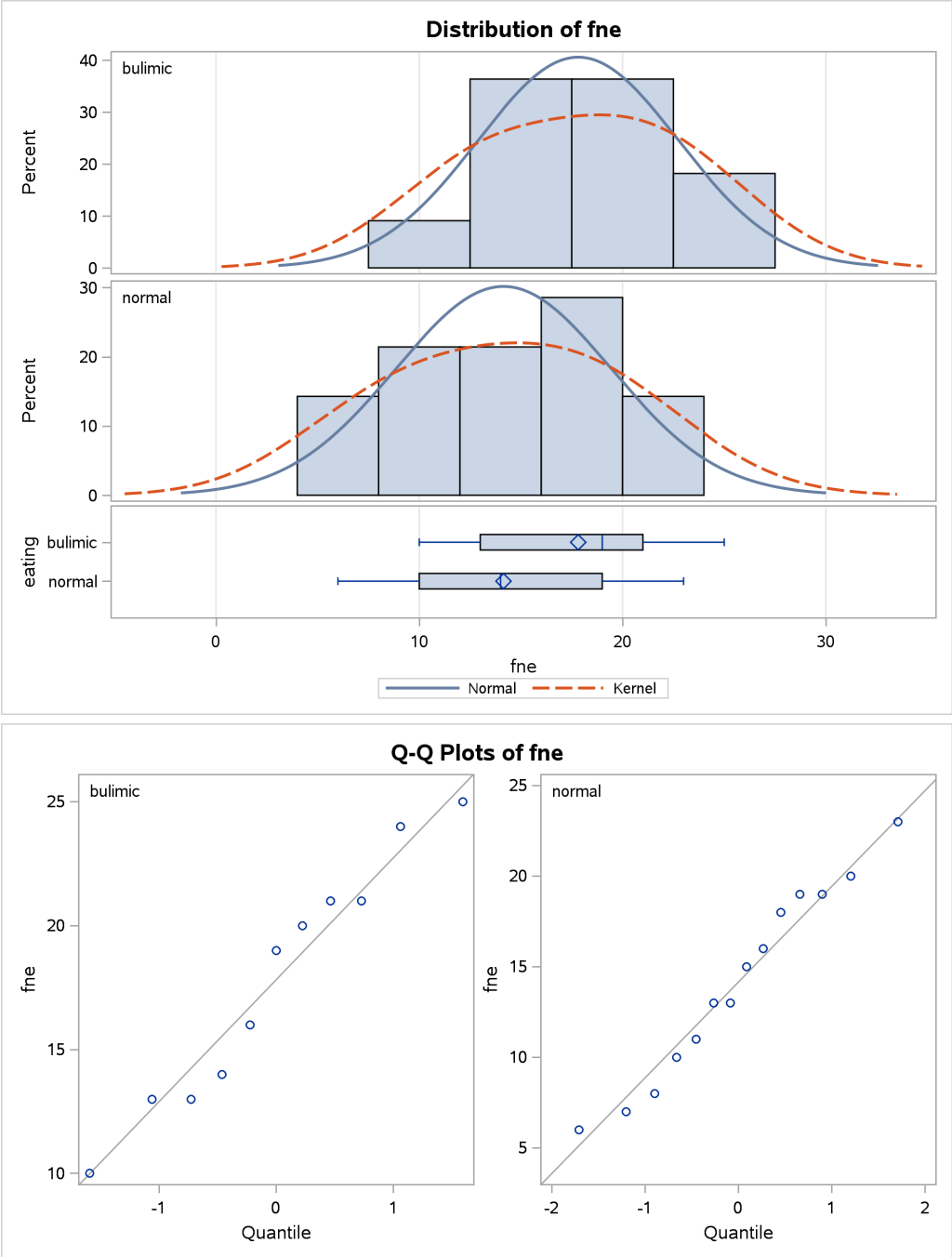

Figure 5: FNE t-test, graphic output

Obs	bulimic	normal
1	21	13
2	13	6
3	10	16
4	20	13
5	25	8
6	19	19
7	16	23
8	21	18
9	24	11
10	13	19
11	14	7
12	.	10
13	.	15
14	.	20

Figure 6: FNE data, alternative layout

Obs	audience	time
1	none	45
2	none	50
3	none	43
4	none	48
5	none	55
6	none	59
7	none	45
8	experimenter	50
9	experimenter	52
10	experimenter	51
11	experimenter	48
12	experimenter	57
13	experimenter	59
14	experimenter	50.5
15	peers	43
16	peers	42
17	peers	47
18	peers	42
19	peers	50
20	peers	39
21	peers	52
22	faculty	41
23	faculty	42
24	faculty	46
25	faculty	49
26	faculty	41
27	faculty	37
28	faculty	50

Figure 7: Stress data

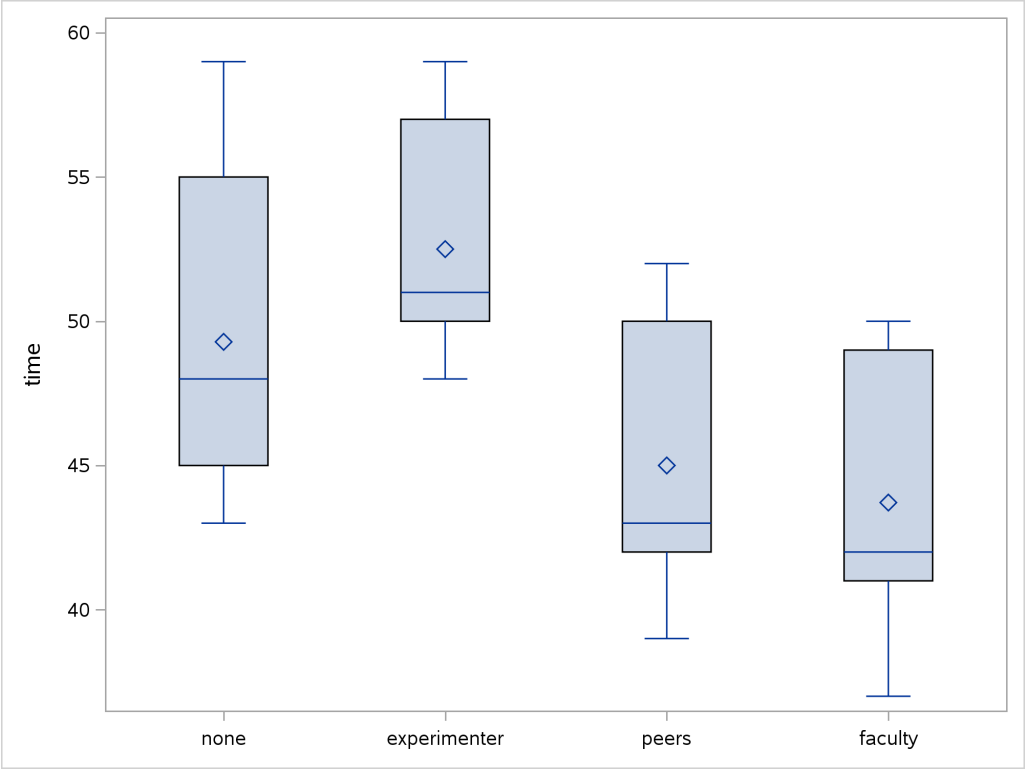

Figure 8: Stress boxplot

```
biking
```

```
## # A tibble: 19 x 3
##    Oxnard  Rose  Rice
##     <dbl> <dbl> <dbl>
##  1    732   869   694
##  2    842   648   629
##  3    736  1045   863
##  4    732   674   748
##  5    736   821   767
##  6    833   708   574
##  7    655   840   628
##  8    688  1029   637
##  9    727   735   620
## 10    721   745   752
## 11    695   794   608
## 12    707   652   983
## 13    843   552   765
## 14    852   732   666
## 15    789   578   727
## 16     NA   661   729
## 17     NA   657   605
## 18     NA   869   717
## 19     NA    NA   679
```

Figure 9: Biking data

```
biking_long
```

```
## # A tibble: 52 x 2
##    street  time
##    <chr>  <dbl>
##  1 Oxnard   732
##  2 Rose     869
##  3 Rice     694
##  4 Oxnard   842
##  5 Rose     648
##  6 Rice     629
##  7 Oxnard   736
##  8 Rose    1045
##  9 Rice     863
## 10 Oxnard   732
## # ... with 42 more rows
```

Figure 10: Biking data, long format

```
uptake
```

```
## # A tibble: 3 x 10
##      id non_bike non_tread non_step mod_bike mod_tread mod_step heavy_bike heavy_tread heavy_step
##   <dbl>    <dbl>     <dbl>    <dbl>    <dbl>     <dbl>    <dbl>      <dbl>       <dbl>      <dbl>
## 1     1     12.8      16.2     22.6     10.9      15.5     20.1        8.7        14.7       16.2
## 2     2     13.5      18.1     19.3     11.1      13.8     21          9.2        13.2       16.1
## 3     3     11.2      17.8     18.9      9.8      16.2     15.9        7.5         8.1       17.8
```

Figure 11: Oxygen uptake data

```
## # A tibble: 27 x 4
##       id smoke exercise oxygen
##    <dbl> <chr> <chr>     <dbl>
##  1     1 non   bike       12.8
##  2     1 non   tread      16.2
##  3     1 non   step       22.6
##  4     1 mod   bike       10.9
##  5     1 mod   tread      15.5
##  6     1 mod   step       20.1
##  7     1 heavy bike        8.7
##  8     1 heavy tread      14.7
##  9     1 heavy step       16.2
## 10     2 non   bike       13.5
## # ... with 17 more rows
```

Figure 12: Oxygen uptake data, tidied, some

```
visits sales
  11     3
  19    11
  16     8
  13     5
  28     8
   5     2
  20     5
  14     6
  22     8
   7     3
  15     5
  29    10
   8     6
  25    10
  16     7
```

Figure 13: Insurance sales data

```
## Parsed with column specification:
## cols(
##  visits = col_double(),
##  sales = col_double()
## )
```

```
ggplot(insurance, aes(x=visits, y=sales)) + geom_point()
```

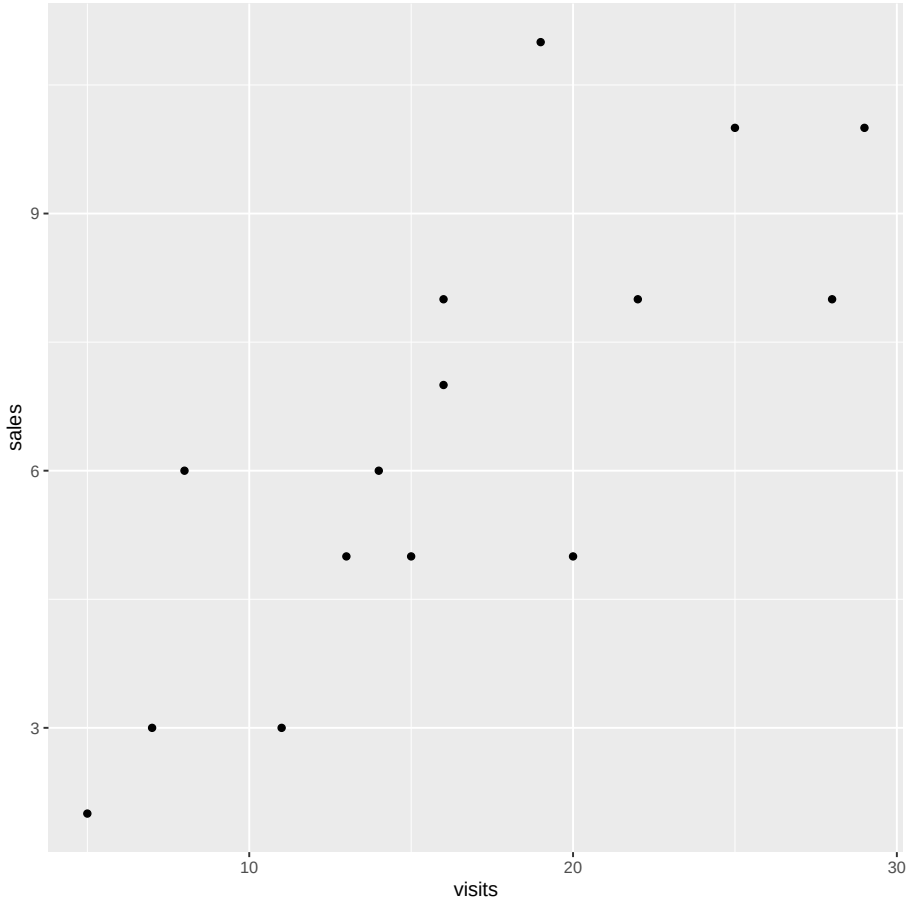

Figure 14: Insurance sales scatterplot

```r
insurance.1 <- lm(sales~visits, data=insurance)
summary(insurance.1)
```

```
##
## Call:
## lm(formula = sales ~ visits, data = insurance)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4817 -1.0538 -0.1167  0.8720  3.8111
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.62597    1.13290   1.435 0.174842
## visits       0.29278    0.06296   4.650 0.000455 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.731 on 13 degrees of freedom
## Multiple R-squared:  0.6245,Adjusted R-squared:  0.5956
## F-statistic: 21.62 on 1 and 13 DF,  p-value: 0.0004545
```

Figure 15: Insurance sales regression

```
ggplot(insurance.1, aes(sample=.resid)) + stat_qq() + stat_qq_line()
```

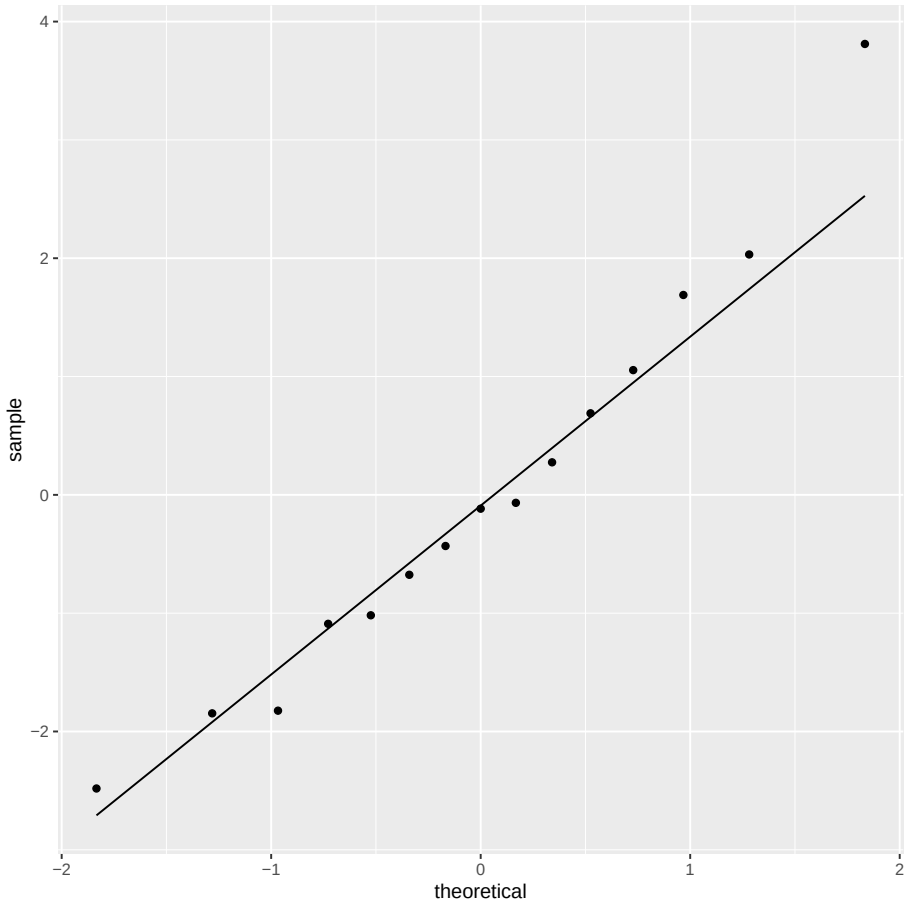

Figure 16: Insurance sales, extra plot 1

```
ggplot(insurance.1, aes(x=.fitted, y=.resid)) + geom_point()
```

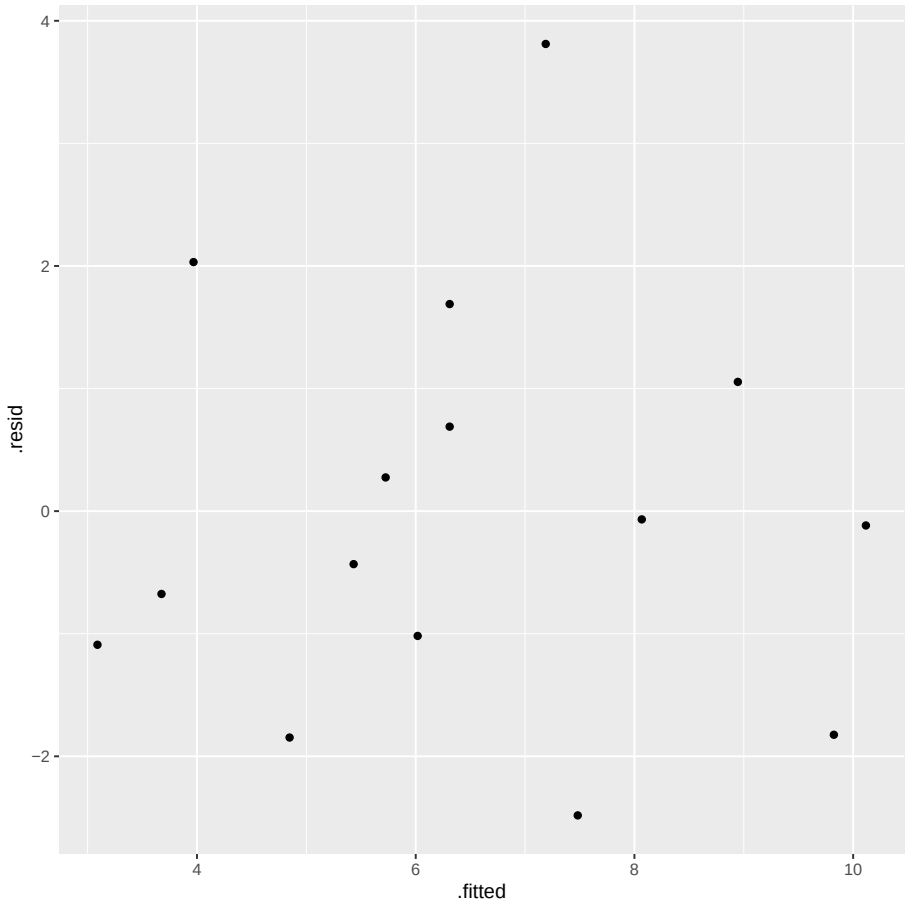

Figure 17: Insurance sales, extra plot 2

```
## # A tibble: 38 x 4
##    country        life_exp tv_per_person dr_per_thousand
##    <chr>             <dbl>         <dbl>           <dbl>
##  1 Argentina          70.5       0.25             2.70
##  2 Bangladesh         53.5       0.00317          0.162
##  3 Brazil             65         0.25             1.46
##  4 Canada             76.5       0.588            2.23
##  5 China              70         0.125            1.56
##  6 Colombia           71         0.179            0.645
##  7 Egypt              60.5       0.0667           1.62
##  8 Ethiopia           51.5       0.00199          0.0273
##  9 France             78         0.385            2.48
## 10 Germany            76         0.385            2.89
## 11 India              57.5       0.0227           0.405
## 12 Indonesia          61         0.0417           0.135
## 13 Iran               64.5       0.0435           0.334
## 14 Italy              78.5       0.263            4.29
## 15 Japan              79         0.556            1.64
## 16 Kenya              61         0.0104           0.131
## 17 Korea, North       70         0.0111           2.70
## 18 Korea, South       70         0.204            0.938
## 19 Mexico             72         0.152            1.67
## 20 Morocco            64.5       0.0476           0.205
## 21 Burma              54.5       0.00169          0.287
## 22 Pakistan           56.5       0.0137           0.423
## 23 Peru               64.5       0.0714           0.984
## 24 Philippines        64.5       0.114            0.942
## 25 Poland             73         0.256            2.08
## 26 Romania            72         0.167            1.79
## 27 Russia             69         0.312            3.86
## 28 South Africa       64         0.0909           0.746
## 29 Spain              78.5       0.385            3.64
## 30 Sudan              53         0.00435          0.0797
## 31 Taiwan             75         0.312            1.04
## 32 Thailand           68.5       0.0909           0.205
## 33 Turkey             70         0.2              0.841
## 34 Ukraine            70.5       0.333            4.42
## 35 United Kingdom     76         0.333            1.64
## 36 United States      75.5       0.769            2.48
## 37 Venezuela          74.5       0.179            1.74
## 38 Vietnam            65         0.0345           0.323
```

Figure 18: Life expectancy data

```r
life.1 <- lm(life_exp~tv_per_person+dr_per_thousand, data=life)
summary(life.1)
```

```
##
## Call:
## lm(formula = life_exp ~ tv_per_person + dr_per_thousand, data = life)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -8.716 -4.131  1.080  3.442  6.468
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      60.1115     1.2226  49.168  < 2e-16 ***
## tv_per_person    24.0169     5.3416   4.496 7.27e-05 ***
## dr_per_thousand   2.0916     0.7902   2.647   0.0121 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.675 on 35 degrees of freedom
## Multiple R-squared:  0.6603,Adjusted R-squared:  0.6409
## F-statistic: 34.02 on 2 and 35 DF,  p-value: 6.21e-09
```

Figure 19: Life expectancy regression 1

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

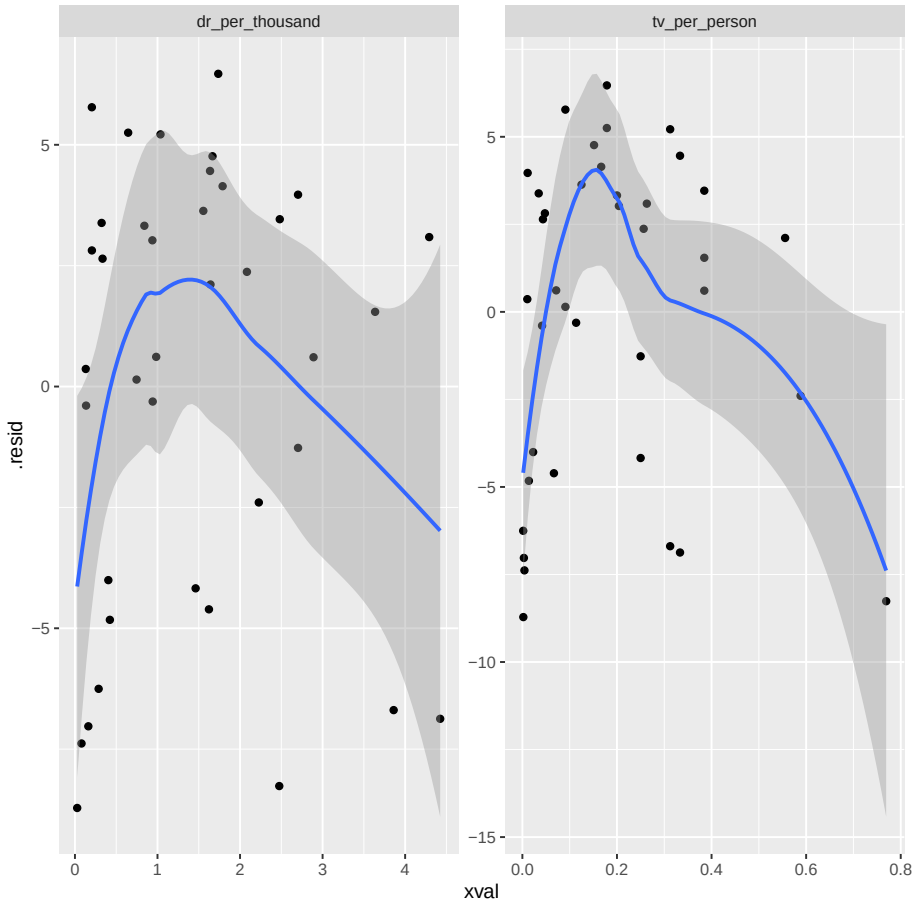

Figure 20: Life expectancy residuals, model 1

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

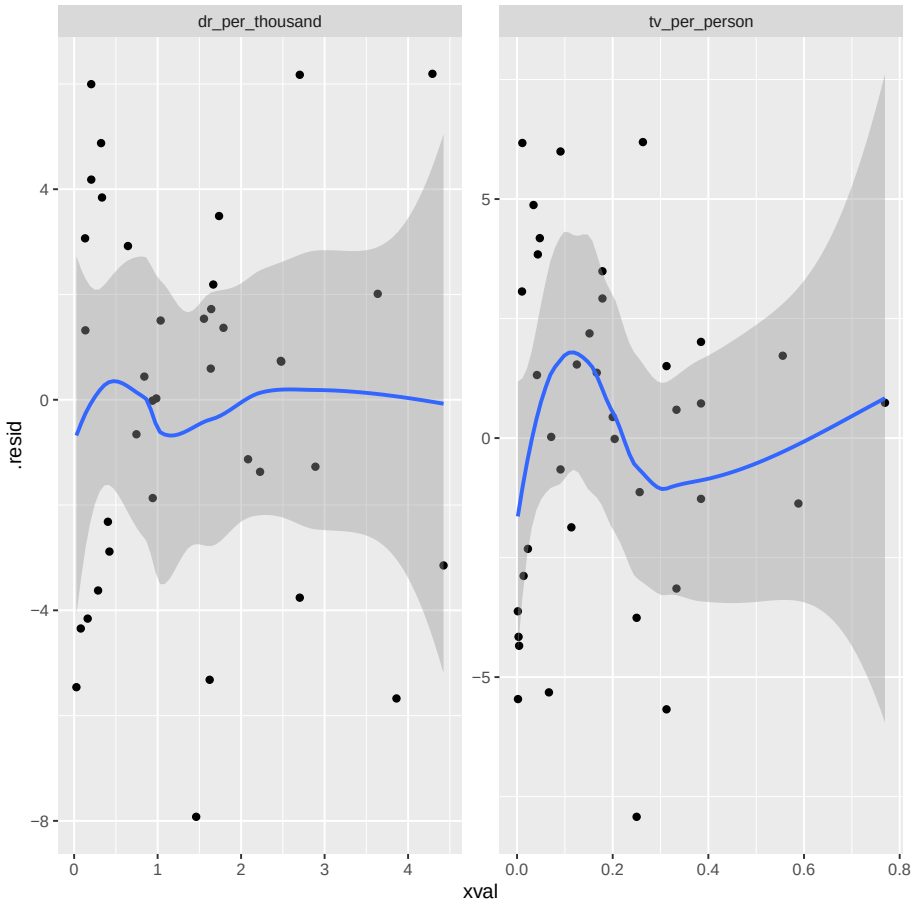

Figure 21: Life expectancy residuals, model 2

```
##
## Call:
## lm(formula = life_exp ~ tv_per_person + dr_per_thousand + I(tv_per_person^2) +
##     I(dr_per_thousand^2), data = life)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.9222 -2.7419  0.5167  2.1453  6.1920
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            56.7129     1.2969  43.729  < 2e-16 ***
## tv_per_person          58.3550    12.6006   4.631 5.45e-05 ***
## dr_per_thousand         4.8297     2.1168   2.282  0.02909 *
## I(tv_per_person^2)    -56.2349    16.2538  -3.460  0.00151 **
## I(dr_per_thousand^2)   -0.9009     0.4367  -2.063  0.04703 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.77 on 33 degrees of freedom
## Multiple R-squared:  0.7918,Adjusted R-squared:  0.7666
## F-statistic: 31.37 on 4 and 33 DF,  p-value: 8.006e-11
```

Figure 22: Life expectancy, alternative model

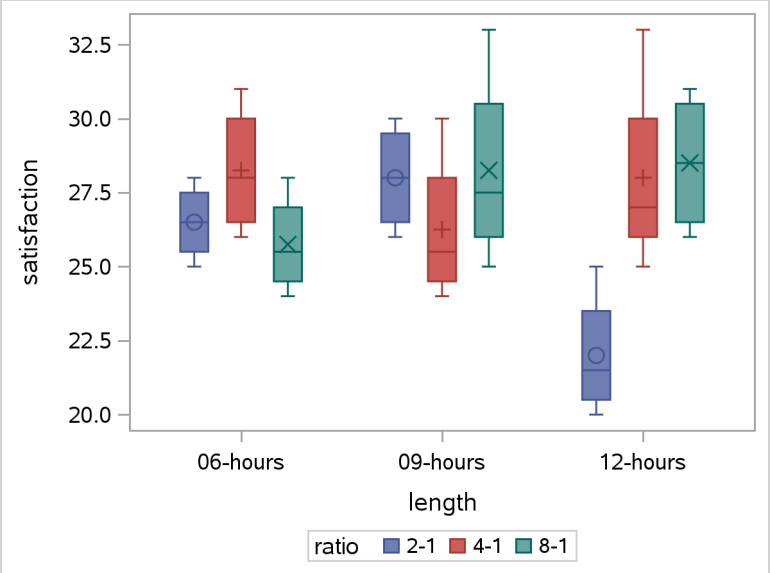

Figure 23: Time-of-day electricity pricing graph