Central Appalachian Mine Closures

Hannah Smith

Contents

List of Tables

List of Figures

1 Rationale and Research Questions

This project seeks to analyze the migration patterns of residents of coal counties in central Appalachia from the years 2000 to 2017. Two "coal counties" serve as a case study for migration in central Appalachia: Harlan County, Kentucky and Dickenson County, Virginia. Harlan County serves as a "Boom Bust" county, meaning the county experienced a coal production boom in 2000 and a coal production bust in 2010. Dickenson, on the other hand, is a "Bust Bust" county, where the county experienced a coal production bust in 2000 and did not recover by 2010.

An important portion of understanding migration in Appalachia is investigating the mine operations within the coal counties. As such, this project creates and analyzes a visualization of mine production and employee in Harlan and Dickenson counties between the years 2000 and 2011. The data analysis of the coal production and the number of employees in coal mines in each county could be indicative of migration patterns within the central Appalachian region.

2 Dataset Information

The data in this repository is coal mine data compiled by the Coal and America Bass Connections team at Duke University from the Energy Information Administration (EIA) online database throughout the fall of 2019. The data in this repository is coal mine data compiled by the Coal and America Bass Connections team at Duke University from the Energy Information Administration (EIA) online database throughout the fall of 2019.

"This report is mandatory under the Federal Energy Administration Act of 1974 (Public Law 93-275). Failure to comply may result in criminal fines, civil penalties, and other sanctions as provided by law. Title 18 USC 1001 makes it a criminal offense for any person knowingly and willingly to make to any Agency or Department of the United States any false, fictitious, or fraudulent statements as to any matter within its jurisdiction.

All coal mining companies that owned a mining operation which produced 25,000 or more short tons of coal during the reporting year must submit form EIA-7A, except for anthracite mines. All anthracite mines that produced 10,000 or more short tons during the reporting year must submit form EIA-7A. Standalone facilities (e.g., preparation plant/tipple/loading dock/train loadout) that worked 5,000 or more hours must submit the EIA-7A. Submit a separate form EIA-7A for each mining operation and standalone facility that meets the reporting criteria.

The U.S. Energy Information Administration's (EIA) Form EIA-7A, Annual Survey of Coal Production and Preparation, collects coal production data from U.S. coal mining companies. This includes information on the type and status of coal operations, characteristics of coalbeds mined, recoverable reserves, productive capacity and the disposition of coal mined which provides Congress with basic statistics concerning coal supply. These data appear in the Annual Coal Report, the Quarterly Coal Report, the Monthly Energy Review, and the Annual Energy Review. In addition, the EIA uses the data for coal supply analyses and in short-term modeling efforts, which produce forecasts of coal supply and prices requested by Congress. The forecast data also appear in the Short-Term Energy Outlook and the Annual Energy Outlook."

Therefore, the data used in this project should be timely and accurate. Furthermore, coal production during the decade this project explores was extremely variable. As such, there should be no exclusion of outliers in this report.

3 Exploratory Analysis

3.1 Data Exploration

Data exploration of the Harlan and Dickenson County raw data files.

```
dim(harlan.raw)
```

```
## [1] 509  15
```

```
dim(dickenson.raw)
```

```
## [1] 194  15
```

```
str(harlan.raw)
```

```
## 'data.frame':    509 obs. of  15 variables:
##  $ year                   : int  2011 2011 2011 2011 2011 2011 2011 2011 2011 2011
##  $ mine.name              : Factor w/ 126 levels "# 1","# 2","# 3",..: 34 104 122 4
##  $ mine.state             : Factor w/ 1 level "Kentucky (East)": 1 1 1 1 1 1 1 1 1 1
##  $ countystr              : Factor w/ 1 level "Harlan": 1 1 1 1 1 1 1 1 1 1 ...
##  $ mine.basin             : Factor w/ 1 level "Appalachia Central": 1 1 1 1 1 1 1 1 1
##  $ mine.status            : Factor w/ 6 levels "Active","Active, men not working, n
##  $ mine.type              : Factor w/ 1 level "Underground": 1 1 1 1 1 1 1 1 1 1 ..
##  $ company.type           : Factor w/ 3 levels "Contractor","Indepedent Producer Op
##  $ operation.type         : Factor w/ 2 levels "Mine only","Preparation Plant": 1 2
##  $ operating.company      : Factor w/ 102 levels "A & M Coal Co Inc",..: 54 64 37 3
##  $ operating.company.address: Factor w/ 92 levels "1160 Jackson Dr, Paris, TN 38242",
##  $ production.stons        : Factor w/ 400 levels "0","1,013,140",..: 248 1 1 1 1 1
##  $ average.employees       : int  110 25 22 20 2 9 127 2 35 4 ...
##  $ labor.hours             : Factor w/ 468 levels "","1,211","1,692",..: 174 243 365
##  $ ARC                     : int  1 1 1 1 1 1 1 1 1 1 ...
```

```
str(dickenson.raw)
```

```
## 'data.frame':    194 obs. of  15 variables:
##  $ year                   : int  2011 2011 2011 2011 2011 2011 2011 2011 2011 2011
##  $ mine.name              : Factor w/ 56 levels "#1","#2","#44",..: 37 16 36 32 11
##  $ mine.state             : Factor w/ 1 level "Virginia": 1 1 1 1 1 1 1 1 1 1 1 ...
##  $ countystr              : Factor w/ 1 level "Dickenson": 1 1 1 1 1 1 1 1 1 1 1 ...
##  $ mine.basin             : Factor w/ 1 level "Appalachia Central": 1 1 1 1 1 1 1 1 1
##  $ mine.status            : Factor w/ 5 levels "Active","Active, men working, not p
##  $ mine.type              : Factor w/ 1 level "Underground": 1 1 1 1 1 1 1 1 1 1 ..
##  $ company.type           : Factor w/ 3 levels "Contractor","Indepedent Producer Op
##  $ operation.type         : Factor w/ 2 levels "Mine only","Preparation Plant": 1 1
##  $ operating.company      : Factor w/ 57 levels "Apple Jacks Coal Company Inc",..:
##  $ operating.company.address: Factor w/ 61 levels "1229 Homecreek Rd., Big Rock, VA 2
##  $ production.stons        : Factor w/ 163 levels "0","1,719","103,817",..: 87 118 7
```

```
##  $ average.employees       : int  16 17 13 14 94 48 37 101 36 7 ...
##  $ labor.hours             : Factor w/ 168 levels "","1,689","101,738",..: 97 127 84
##  $ ARC                     : int  1 1 1 1 1 1 1 1 1 1 ...
```

colnames(harlan.raw)

```
##  [1] "year"                   "mine.name"
##  [3] "mine.state"             "countystr"
##  [5] "mine.basin"             "mine.status"
##  [7] "mine.type"              "company.type"
##  [9] "operation.type"         "operating.company"
## [11] "operating.company.address" "production.stons"
## [13] "average.employees"      "labor.hours"
## [15] "ARC"
```

colnames(dickenson.raw)

```
##  [1] "year"                   "mine.name"
##  [3] "mine.state"             "countystr"
##  [5] "mine.basin"             "mine.status"
##  [7] "mine.type"              "company.type"
##  [9] "operation.type"         "operating.company"
## [11] "operating.company.address" "production.stons"
## [13] "average.employees"      "labor.hours"
## [15] "ARC"
```

summary(harlan.raw)

```
##       year             mine.name            mine.state    countystr
##  Min.   :2000   Mine #1        : 24   Kentucky (East):509   Harlan:509
##  1st Qu.:2003   Prep Plant     : 21
##  Median :2006   Mine No 1      : 19
##  Mean   :2006   Mine No. 1     : 11
##  3rd Qu.:2008   No 1 Plant     : 11
##  Max.   :2011   Darby Fork No 1: 10
##                 (Other)        :413
##            mine.basin                               mine.status
##  Appalachia Central:509   Active                            :420
##                           Active, men not working, not producing:  1
##                           Active, men working, not producing   : 34
##                           Mine closed by MSHA                   : 21
##                           Permanently abandoned                 : 27
##                           Temporarily closed                    :  6
##
##        mine.type                       company.type         operation.type
##  Underground:509   Contractor                  : 92   Mine only        :399
##                    Indepedent Producer Operator:324   Preparation Plant:110
```

```
##                    Operating Subsidiary       : 93
##
##
##
##
##                         operating.company
##   Harlan Cumberland Coal Company: 46
##   Lone Mountain Processing Inc  : 31
##   Manalapan Mining Co., Inc.    : 29
##   Manalapan Mining Company Inc  : 28
##   Liggett Mining Llc            : 23
##   Rex Coal Company, Inc.        : 14
##   (Other)                       :338
##                          operating.company.address  production.stons
##   P.O. Box 269, Grays Knob, KY 40829    : 42        0          :110
##   P.O. Box 527, Benham, KY 40807        : 35        1,013,140:  1
##   8174 E Hwy 72, Pathfork, KY 40863     : 31        1,046,339:  1
##   P.O. Box 838, Middlesboro, KY 40965   : 28        1,047,698:  1
##   P.O. Box 1226, Norton, VA 24273       : 25        1,122,057:  1
##   General Delivery, Grays Knob, KY 40829: 22        1,186,166:  1
##   (Other)                               :326        (Other)  :394
##   average.employees  labor.hours       ARC
##   Min.   :  1.00            : 40   Min.   :1
##   1st Qu.: 15.00    23,520 :  3   1st Qu.:1
##   Median : 25.00    1,211  :  1   Median :1
##   Mean   : 34.83    1,692  :  1   Mean   :1
##   3rd Qu.: 40.00    1,702  :  1   3rd Qu.:1
##   Max.   :166.00    1,705  :  1   Max.   :1
##   NA's   :40        (Other):462
```

summary(dickenson.raw)

```
##      year                mine.name        mine.state       countystr
##   Min.   :2000   Cherokee Mine      : 12  Virginia:194   Dickenson:194
##   1st Qu.:2001   Mc Clure River  Plant: 10
##   Median :2004   Nme                : 8
##   Mean   :2005   No 8               : 7
##   3rd Qu.:2007   No. 3 Mine         : 7
##   Max.   :2011   No. 4              : 7
##                  (Other)            :143
##           mine.basin                              mine.status
##   Appalachia Central:194   Active                          :151
##                            Active, men working, not producing: 14
##                            Mine closed by MSHA             :  4
##                            Permanently abandoned           : 21
```

```
##                            Temporarily closed              :  4
##
##
##        mine.type                      company.type          operation.type
##   Underground:194   Contractor                 :50   Mine only       :162
##                     Indepedent Producer Operator:66   Preparation Plant: 32
##                     Operating Subsidiary        :78
##
##
##
##
##                    operating.company
##   DickensonRussell Coal Company : 19
##   Clinchfield Coal Company      : 12
##   DickensonRussell Coal Co Llc  :  9
##   L & J Equipment Company       :  9
##   Apple Jacks Coal Company, Inc.:  7
##   DickensonRussell Coal Co., Ll :  7
##   (Other)                       :131
##                    operating.company.address production.stons
##   P.O. Box 1426, Grundy, VA 24614      : 15    0        : 32
##   5703 Crutchfield Dr, Norton, VA 24273: 11    1,719   :  1
##   P.O. Box 1025, Grundy, VA 24614      :  9    103,817:  1
##   P.O. Box 148, Vansant, VA 24656      :  8    104,108:  1
##   P.O. Box 458, Big Rock, VA 24603     :  8    104,515:  1
##   Rt 2, Box 73, Cleveland, VA 24225    :  8    109,076:  1
##   (Other)                              :135    (Other):157
##   average.employees   labor.hours        ARC
##   Min.   :  1.00              : 26   Min.   :1
##   1st Qu.: 12.00    79,091 :  2   1st Qu.:1
##   Median : 16.00    1,689  :  1   Median :1
##   Mean   : 30.11    101,738:  1   Mean   :1
##   3rd Qu.: 40.00    104,181:  1   3rd Qu.:1
##   Max.   :117.00    107,901:  1   Max.   :1
##   NA's   :26        (Other):162
```

```r
sum(as.numeric(harlan.raw$average.employees), na.rm = TRUE)
```

```
## [1] 16336
```

```r
sum(as.numeric(dickenson.raw$average.employees), na.rm = TRUE)
```

```
## [1] 5058
```

```r
sum(as.numeric(harlan.raw$production.stons), na.rm = TRUE)
```

```
## [1] 80309
```

```r
sum(as.numeric(dickenson.raw$production.stons), na.rm = TRUE)
```

[1] 13397

Formatting Relevant Columns in Harlan and Dickenson County Raw Datasets

```r
harlan.raw$year <- as.numeric(harlan.raw$year)
harlan.raw$ARC <- as.logical(harlan.raw$ARC)
harlan.raw$mine.status <- as.character(harlan.raw$mine.status)
harlan.raw$operation.type <- as.character(harlan.raw$operation.type)
harlan.raw$production.stons <- as.numeric(harlan.raw$production.stons)
harlan.raw$average.employees <- as.integer(harlan.raw$average.employees)


dickenson.raw$year <- as.numeric(dickenson.raw$year)
dickenson.raw$ARC <- as.logical(dickenson.raw$ARC)
dickenson.raw$mine.status <- as.character(dickenson.raw$mine.status)
dickenson.raw$operation.type <- as.character(dickenson.raw$operation.type)
dickenson.raw$production.stons <- as.numeric(dickenson.raw$production.stons)
dickenson.raw$average.employees <- as.numeric(dickenson.raw$average.employees)
```

Visual Data Exploration of Harlan and Dickenson raw data

##Data Wrangling Joining Harlan and Dickenson raw data together

```r
minedata <- full_join(harlan.raw, dickenson.raw)
```

Joining, by = c("year", "mine.name", "mine.state", "countystr",
"mine.basin", "mine.status", "mine.type", "company.type", "operation.type",
"operating.company", "operating.company.address", "production.stons",
"average.employees", "labor.hours", "ARC")

Warning: Column `mine.name` joining factors with different levels, coercing to
character vector

Warning: Column `mine.state` joining factors with different levels, coercing to
character vector

Warning: Column `countystr` joining factors with different levels, coercing to
character vector

Warning: Column `operating.company` joining factors with different levels,
coercing to character vector

Warning: Column `operating.company.address` joining factors with different
levels, coercing to character vector

Warning: Column `labor.hours` joining factors with different levels, coercing to
character vector

Filtering for data relevant to research questions

```r
harlan.clean <- harlan.raw %>%
  select(year:countystr, mine.status, operation.type, operating.company,
         production.stons:labor.hours) %>%
  filter(mine.status == "Active", operation.type == "Mine only") %>%
  group_by(year, countystr) %>%
  mutate(total.stons = sum(production.stons), total.employee = sum(average.employees))

dickenson.clean <- dickenson.raw %>%
  select(year:countystr, mine.status, operation.type, operating.company,
         production.stons:labor.hours) %>%
  filter(mine.status == "Active", operation.type == "Mine only") %>%
  group_by(year, countystr) %>%
  mutate(total.stons = sum(production.stons), total.employee = sum(average.employees))

minedata.clean <- minedata %>%
  select(year:countystr, mine.status, operation.type, operating.company,
         production.stons:labor.hours) %>%
  filter(mine.status == "Active", operation.type == "Mine only") %>%
  group_by(year, countystr) %>%
  mutate(total.stons = sum(production.stons), total.employee = sum(average.employees))
```

Save the data in the processed folder

```r
write.csv(harlan.clean, "./Data/Processed/EIA_MineData_Harlan_Processed.csv")
write.csv(dickenson.clean, "./Data/Processed/EIA_MineData_Dickenson_Processed.csv")
write.csv(minedata.clean, "./Data/Processed/EIA_MineData_Combined_Processed.csv")
```

4 Analysis

4.1 Question 1: Is the annual tonnage produced by all mines in a boom-bust (Harlan) county in the years 2000 to 2011 significantly greater than that produced by a bust-bust (Dickenson) county in the year 2000 to 2011?

```
## Scale for 'x' is already present. Adding another scale for 'x', which will
## replace the existing scale.
```
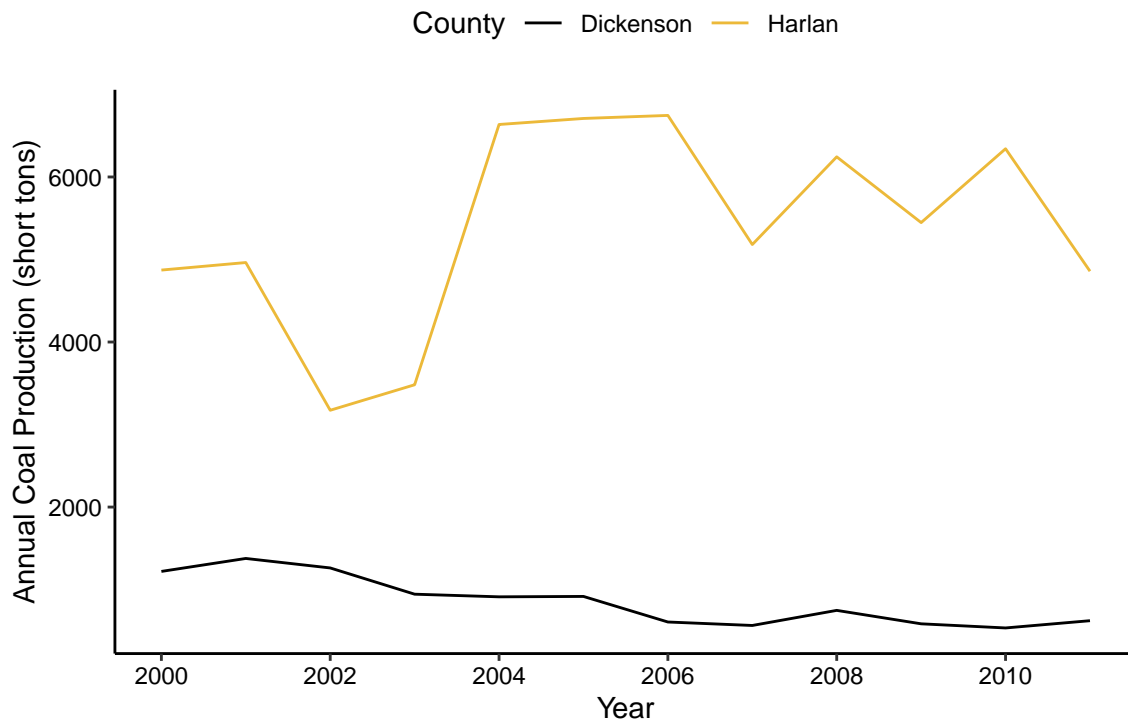

Figure 1: Total Annual Coal Production in Harlan and Dickenson Counties

4.2 Question 2: Is the number of people employed in all mines in the boom-bust county (Harlan) in the years 2001 to 2011 significantly greater than those employed by a bust-bust county (Dickenson) in the years 2001 to 2011?

```
employee <- ggplot(minedata.clean, aes(x = year, y = total.employee, color = countystr))
  geom_line() +
  labs(x = "Year", y = "Total Employed", color = "County") +
  xlim(2001, 2011) +
  scale_color_manual(values=hppal2) +
```

```
   scale_x_continuous("Year", c(2000, 2002, 2004, 2006, 2008, 2010))
```

```
## Scale for 'x' is already present. Adding another scale for 'x', which will
## replace the existing scale.
```

```
print(employee)
```

```
## Warning: Removed 43 rows containing missing values (geom_path).
```
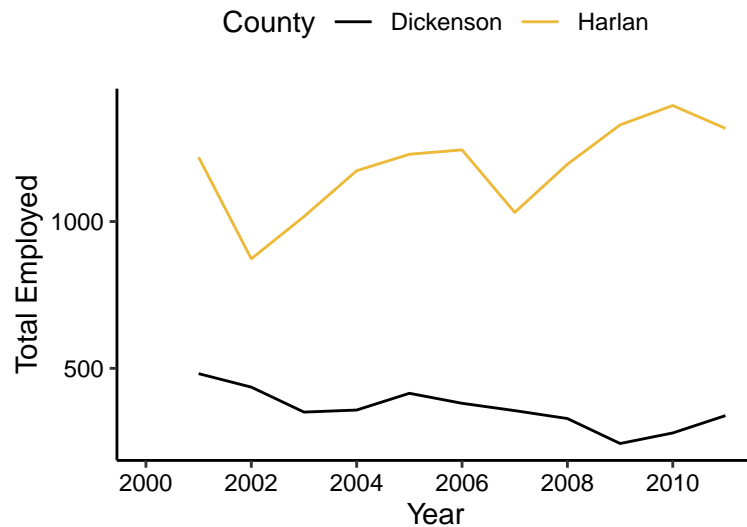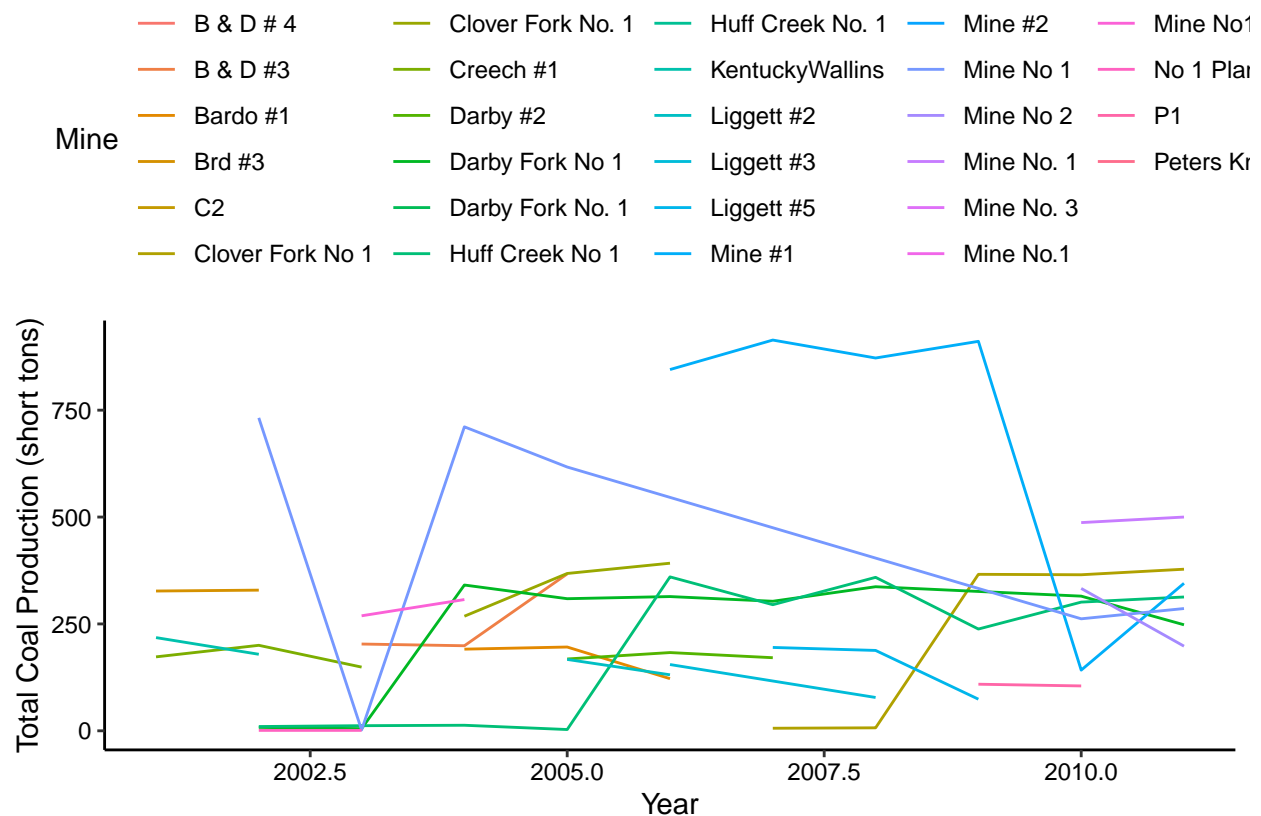


```
ggsave("./Output/AnnualEmployment.jpg", employee, height = 4, width = 6, dpi = 300)
```

```
## Warning: Removed 43 rows containing missing values (geom_path).
```

Harlan Mines by Employee

Harlan Mines by Production

```
harlan.production.plot <- ggplot(harlan.clean.bymine, aes(x = year, y = total.stons, col
 geom_line() +
  labs(x = "Year", y = "Total Coal Production (short tons)", color = "Mine")
  #lim(2001, 2011) +
  #scale_color_manual(values=hppal31)
print(harlan.production.plot)
```

Legend (Mine):
B & D # 4, B & D #3, Bardo #1, Brd #3, C2, Clover Fork No 1, Clover Fork No. 1, Creech #1, Darby #2, Darby Fork No 1, Darby Fork No. 1, Huff Creek No 1, Huff Creek No. 1, KentuckyWallins, Liggett #2, Liggett #3, Liggett #5, Mine #1, Mine #2, Mine No 1, Mine No 2, Mine No. 1, Mine No. 3, Mine No.1, Mine No1, No 1 Pla, P1, Peters Kn

Dickenson Mines by Employee

```
dickenson.clean.bymine <- dickenson.raw %>%
  filter(average.employees > 50) %>%
  group_by(year, mine.name) %>%
  mutate(total.stons = sum(production.stons), total.employee = sum(average.employees))

dickenson.employee.plot <- ggplot(dickenson.clean.bymine, aes(x = year, y = total.employ
 geom_line() +
  labs(x = "Year", y = "Number of Employees", color = "Mine") +
  xlim(2001, 2011) +
  scale_color_manual(values=hppal33) +
  scale_x_continuous("Year", c(2000, 2002, 2004, 2006, 2008, 2010))
```

```
## Scale for 'x' is already present. Adding another scale for 'x', which will
## replace the existing scale.
```

```
  #lim(2001, 2011) +
  #scale_color_manual(values=hppal31)
print(dickenson.employee.plot)
```
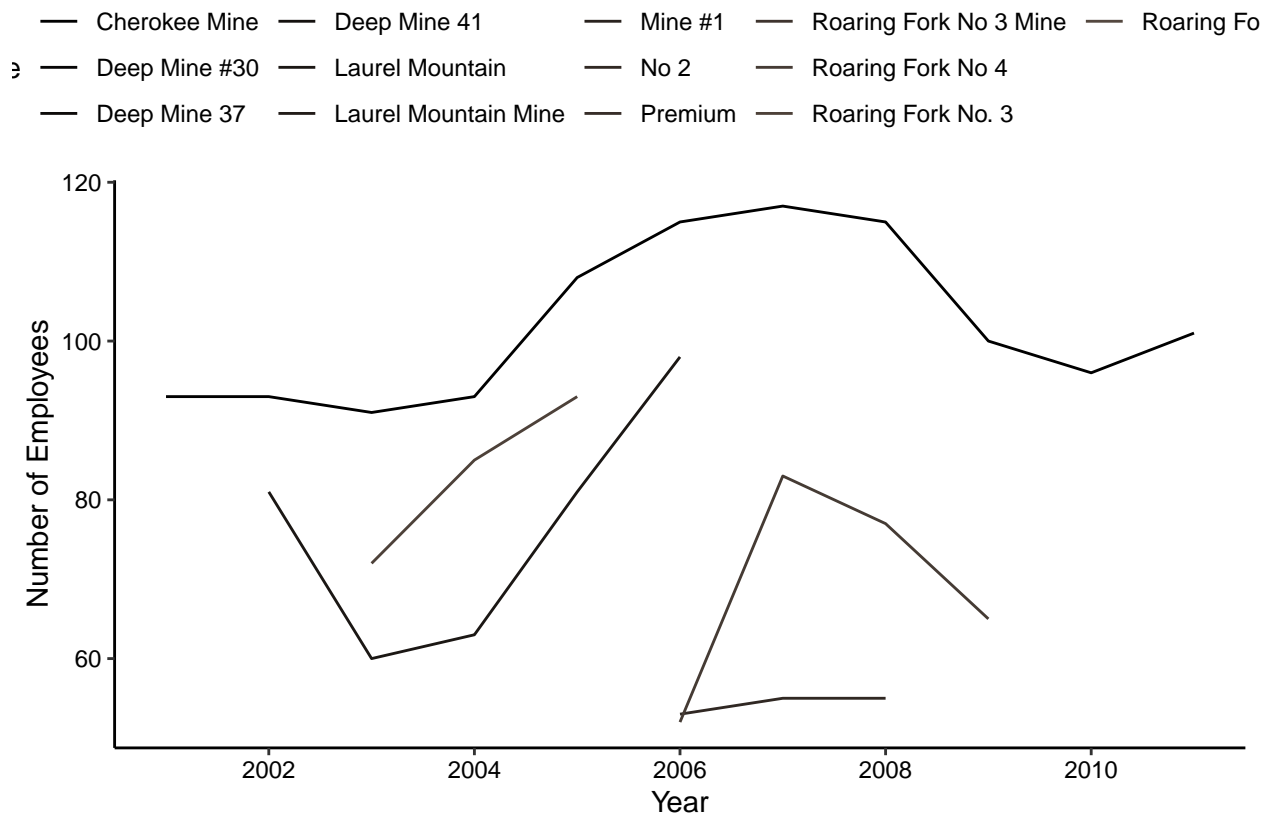
Dickenson Mines by Production

```
dickenson.production.plot <- ggplot(dickenson.clean.bymine, aes(x = year, y = total.ston
 geom_line() +
  labs(x = "Year", y = "Total Coal Production (short tons)", color = "Mine")
  #lim(2001, 2011) +
  #scale_color_manual(values=hppal31)
print(dickenson.production.plot)
```
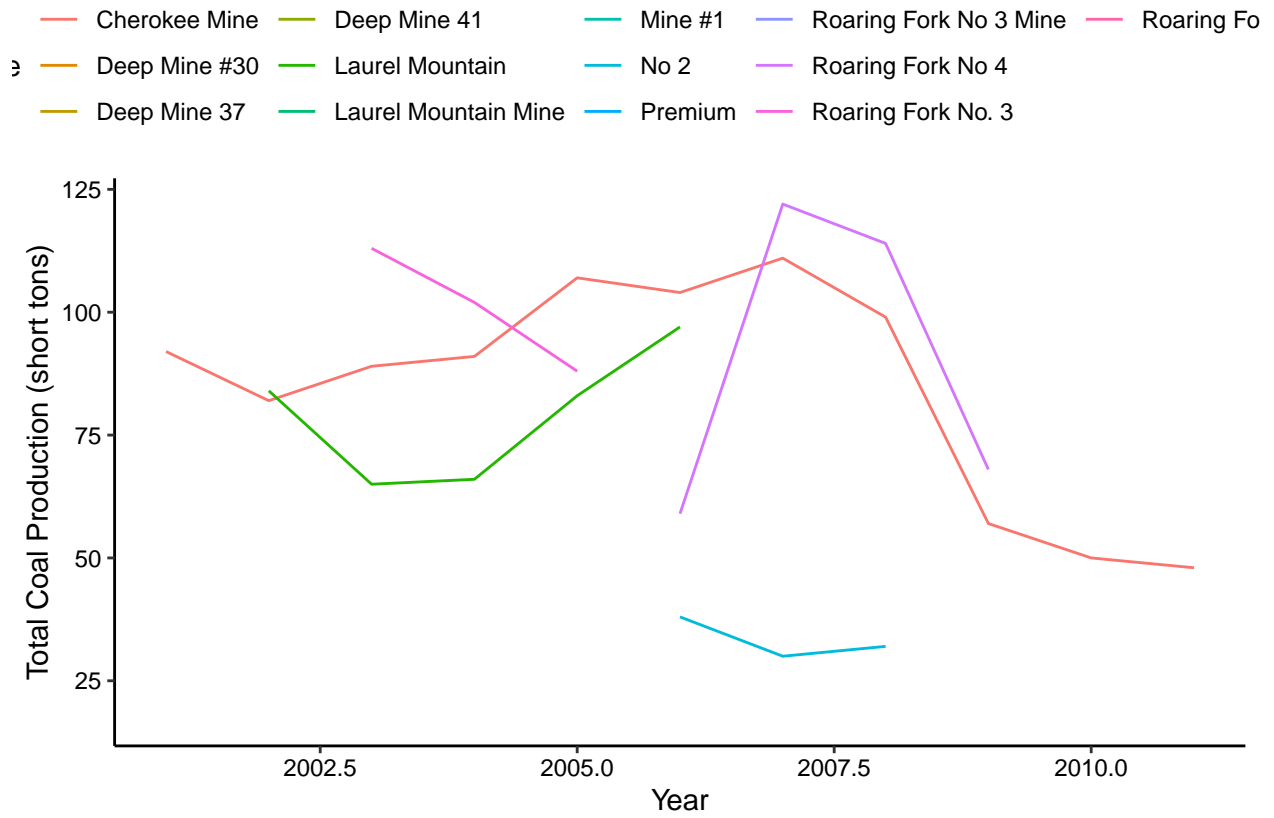
5 Summary and Conclusions

6   References