Linear and k-means modeling

Àlex Fabregat

Pick a commodity.

It is necessary to generate a bunch of local commodity csv files.

Generate the data frame with all the node variables.

```
source("model_get_data.R")
all_info <- model_get_data(route)
```

Do the k-means

Pick a country and it'll return an ordered list of the "most similar" countries.

```
source("model_kmeans.R")
km <- model_kmeans(all_info, "United Kingdom")
print(km[1:5,])
```

```
##      Var1 Freq
## 1 Belgium   11
## 2 Austria   10
## 3 Denmark    8
## 4  Poland    8
## 5  France    6
```

Fit a linear model

```
source("model_linear.R")
lm.fit <- model_linear(all_info)
```

```
## Reordering variables and trying again:
```

```
#summary(lm.fit)
```

Now test the model against a (country, period) pair

Let me cheat a bit here and pick a really good pair :)

```
## [1] "Observed, predicted and error for overall trade value in billions $US: 0.00673, 0.00666, 1.16%"
```

Now, let's show a bad one :(

```
## [1] "Observed, predicted and error for overall trade value in billions $US: 0.01575, 0.00413, 281.25%"
```