
ISyE 2028 – Basic Statistical Methods - Fall 2015

Bonus Project: “Big” Data Analytics Proposal

Methods to Evaluate Each Month’s and Year’s Exchange Rate Change (KRW-USD)

Daniel Kim

Background:

Exchange rate is one of the important factors that give a huge impact on economy. Levels and fluctuations in the exchange rate exert a powerful impact on a nation’s exports, imports and the trade balance. These subjects of exchange rate’s influences, then, impacts on businesses and households in a nation. Even though the daily change in exchange rate highly depends on many factors that happen irregularly and unexpectedly, as a student whose parents live in South Korea, I would like to investigate in which month the exchange rate most likely increases or decreases, when the exchange rate between two countries was the highest and lowest, and see if there is a trend or pattern in the exchange rate.

Method:

Using python programming language and exchange rate website, x-rates.com, I collected average exchange rate for each month from 2005 January to 2015 November. These samples are not clearly biased.

```
import urllib.request
import csv
def isye2028project():
    year = 5
    data1t = []
    while year <= 9: #From Year 2005 to 2009
        yearstr= str(year)
        url = "http://www.x-rates.com/average/?from=USD&to=KRW&amount=1&year=200" + yearstr
        request = urllib.request.Request(url)
        response = urllib.request.urlopen(request)
        page = response.read()
        text = page.decode()

        data = text.find("drawGraph")
        dataBeg = data + 26
        dataend = 0
        while text[dataBeg + dataend] == "|":
            dataend += 1
        dataEnd = dataBeg+dataend

        data1t.append(("Year 200" + yearstr + " : " + text[dataBeg:dataEnd]))
        year += 1

    year = 10
    while year <= 15: #From Year 2010 to 2015
        yearstr= str(year)
        url = "http://www.x-rates.com/average/?from=USD&to=KRW&amount=1&year=20" + yearstr
        request = urllib.request.Request(url)
        response = urllib.request.urlopen(request)
        page = response.read()
        text = page.decode()

        data = text.find("drawGraph")
        dataBeg = data + 26
        dataend = 0
        while text[dataBeg+ dataend] == "|":
            dataend+=1
        dataEnd = dataBeg + dataend

        data1t.append(("Year 20" + yearstr + " : " + text[dataBeg:dataEnd]))
        year += 1

    #return data1t

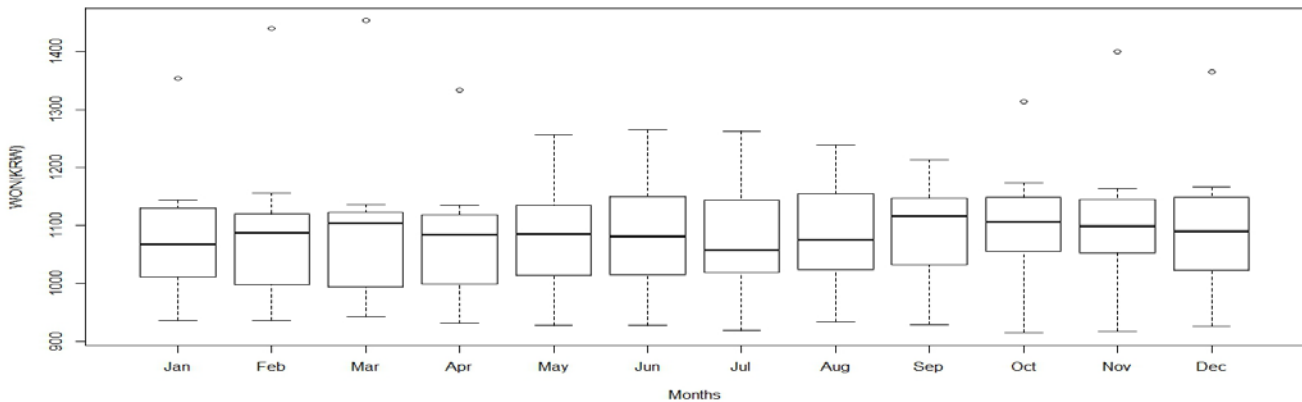
    year1t=["Year 2005", "Year 2006", "Year 2007", "Year 2008", "Year 2009",
            "Year 2010", "Year 2011", "Year 2012", "Year 2013", "Year 2014",
            "Year 2015"]
    introlt = ["Year", "Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec"]

    afile = open("isye2028projects.csv", "w", newline='')
    finallt = []
    finallt.append(introlt)
    csvWriter = csv.writer(afile)
    count = 12
    for index in range(11):
        list1 = []
        list1.append(year1t[index])
        for index2 in range(12):
            list1.append(data1t[index][count:count+15])
            count += 16
        finallt.append(list1)
    csvWriter.writerows(finallt)
    afile.close()
```

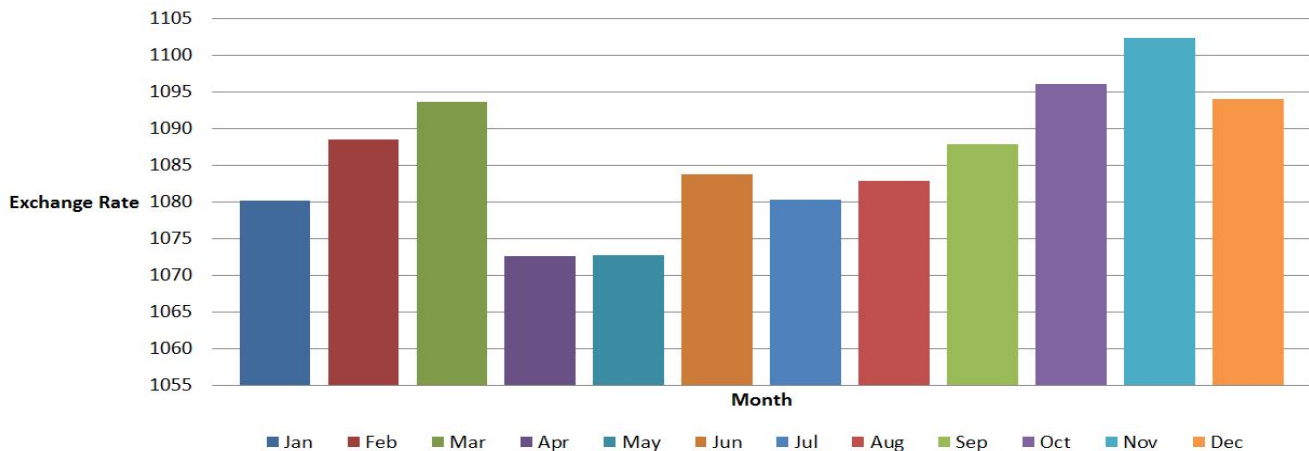
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Avg
2	2005	1039.417363	1023.24086	1007.57357	1009.58758	1001.934558	1011.504546	1036.71342	1022.43942	1029.22896	1046.58569	1041.55582	1022.45097	1024.353
3	2006	984.289351	971.089688	974.29945	953.90746	941.744735	954.361825	950.682581	961.668622	953.235658	953.14687	935.519394	926.427549	955.0311
4	2007	936.93986	937.03991	943.315672	931.009075	927.793541	927.790735	918.475737	934.58614	928.5041	914.839847	917.275163	931.35621	929.0772
5	2008	942.54961	944.497574	981.317117	987.416684	1035.344445	1032.633231	1017.47591	1047.87697	1133.37592	1313.6645	1398.80276	1364.48423	1099.953
6	2009	1353.731521	1438.82173	1452.46298	1332.98537	1256.492359	1264.027019	1261.62589	1238.83268	1213.4305	1173.84049	1163.92144	1165.75667	1276.327
7	2010	1141.417817	1156.57178	1136.26699	1116.05341	1163.683471	1215.259156	1204.52806	1180.55473	1162.33325	1122.43075	1130.35283	1144.84142	1156.191
8	2011	1118.877656	1118.31316	1119.96817	1083.86481	1084.43257	1081.228632	1057.83176	1074.83063	1116.24008	1151.87533	1134.52709	1148.04749	1107.503
9	2012	1143.000722	1122.50421	1125.95972	1134.94042	1157.938804	1164.804938	1142.5805	1132.60804	1123.35533	1106.0457	1087.92834	1075.69205	1126.447
10	2013	1066.030398	1087.80544	1103.74476	1121.66204	1112.627231	1135.877065	1126.35236	1115.91798	1084.14299	1066.69234	1062.75818	1056.55721	1095.014
11	2014	1066.96433	1071.43117	1070.65756	1042.46955	1025.166427	1019.125784	1021.3228	1024.82695	1036.28463	1062.34998	1098.34259	1104.58447	1053.627
12	2015	1088.778397	1102.21161	1114.22336	1084.61329	1093.321198	1114.145066	1145.60694	1177.90883	1186.44646	1144.74808	1155.32004		1127.938
13	AVG	1080.181548	1088.50247	1093.61721	1072.59179	1072.770849	1083.705272	1080.29054	1082.91372	1087.87072	1096.01996	1102.39124	1094.01983	

From this data, I drew two histograms using Excel and two boxplots using R software. From the histogram of average exchange rate for each month, I could see that there is no definite trend of increase or decrease in exchange rate in an year. The exchange rate was the lowest in April or May and highest in November. Even though, histogram shows that there is no definite trend, from the boxplot of average exchange rate for each month from 2005 to 2015, each month's average exchange rate falls in an interval between 1070 and 1105 KRW.

Boxplot of Exchange Rate For Each Month From 2005-2015



Average Exchange Rate For Each Month From 2005-2015



```
> summary(datareal)
```

Jan		Feb		Mar		Apr	
Min.	: 936.9	Min.	: 937.0	Min.	: 943.3	Min.	: 931.0
1st Qu.	:1011.9	1st Qu.	: 997.2	1st Qu.	: 994.4	1st Qu.	: 998.5
Median	:1067.0	Median	:1087.8	Median	:1103.7	Median	:1083.9
Mean	:1090.2	Mean	:1088.5	Mean	:1093.6	Mean	:1072.6
3rd Qu.	:1130.1	3rd Qu.	:1120.4	3rd Qu.	:1123.0	3rd Qu.	:1118.9
Max.	:1353.7	Max.	:1438.8	Max.	:1452.5	Max.	:1333.0

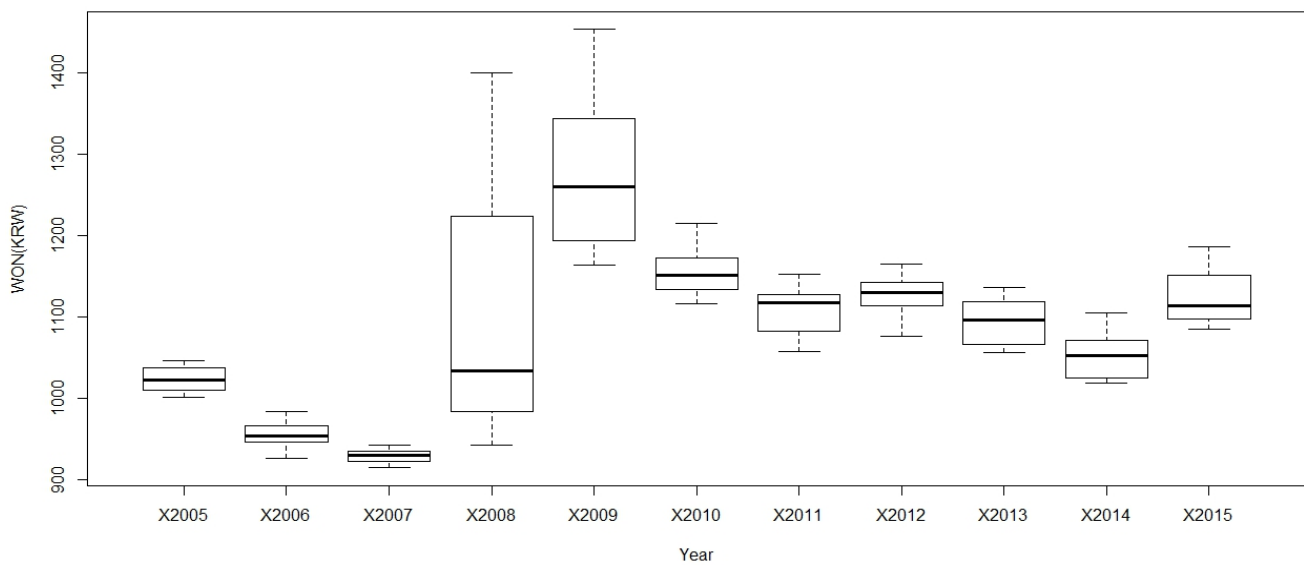
May		Jun		Jul		Aug	
Min.	: 927.8	Min.	: 927.8	Min.	: 918.5	Min.	: 934.6
1st Qu.	:1013.6	1st Qu.	:1015.3	1st Qu.	:1019.4	1st Qu.	:1023.6
Median	:1084.4	Median	:1081.2	Median	:1057.8	Median	:1074.8
Mean	:1072.8	Mean	:1083.7	Mean	:1080.3	Mean	:1082.9
3rd Qu.	:1135.3	3rd Qu.	:1150.3	3rd Qu.	:1144.1	3rd Qu.	:1155.3
Max.	:1256.5	Max.	:1264.0	Max.	:1261.6	Max.	:1238.8

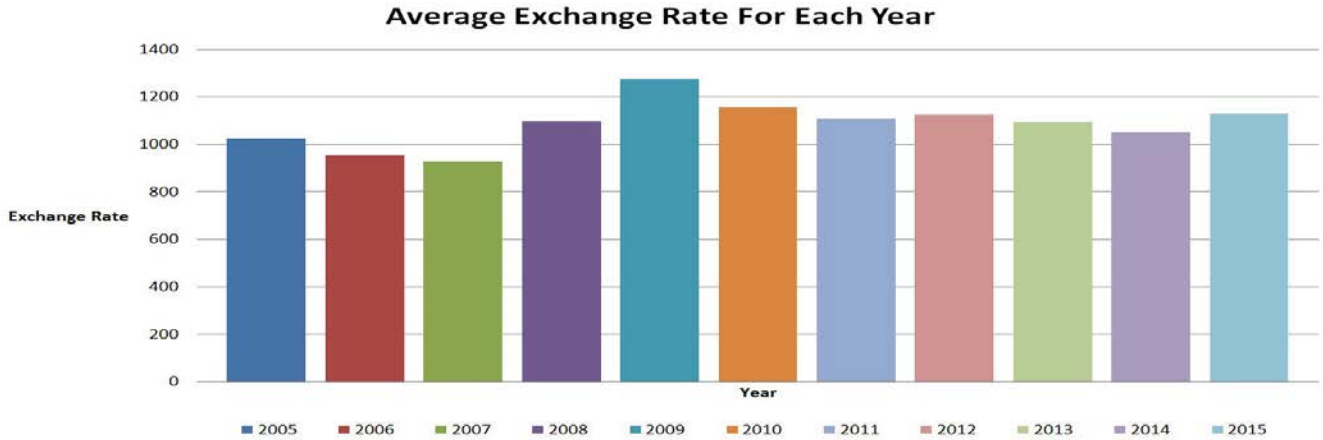
Sep		Oct		Nov		Dec	
Min.	: 928.5	Min.	: 914.8	Min.	: 917.3	Min.	: 926.4
1st Qu.	:1032.8	1st Qu.	:1054.5	1st Qu.	:1052.2	1st Qu.	:1031.0
Median	:1116.2	Median	:1106.0	Median	:1098.3	Median	:1090.1
Mean	:1087.9	Mean	:1096.0	Mean	:1102.4	Mean	:1094.0
3rd Qu.	:1147.9	3rd Qu.	:1148.3	3rd Qu.	:1144.9	3rd Qu.	:1147.2
Max.	:1213.4	Max.	:1313.7	Max.	:1398.8	Max.	:1364.5

The minimum number of exchange rate between South Korea and USA was 914.8 KRW for 1 USD, which occurred in October, 2007. The maximum number of exchange rate was 1398.8 KRW for 1 USD, which occurred in November, 2008.

From the histogram of average exchange rate for each year, I see that the highest average exchange rate between South Korea and USA occurred in 2009 and the lowest average exchange rate occurred in 2007. From this information, as 5-numerical summary for each month shows, I can conclude that the minimum number of exchange rate was not an extreme outlier but the maximum number of exchange rate, which occurred in 2008, was an extreme outlier since it did not happen in 2009. From the boxplot of average exchange rate for each year, the range of boxplot in 2008 was the highest. This means that standard deviation was large and there were many exchange rate changes and fluctuations in 2008. Sample standard deviation for year 2008 is, in fact, 165.1388 KRW. On the other hand, the range and sample standard deviation were low in 2007. Both boxplot and histogram show that after the exchange rate gradually increased from 2005 to 2009, it started decreasing until 2014 and slightly increased again in 2015.

Boxplot of Exchange Rate For Each Year





I also conducted confidence interval method so that anyone can easily find out an outlier in the future. I calculated 95% and 99% confidence interval using t-distribution since population mean and population variance are both unknown.

	A	B	C
1	CI	95%	99%
2	Jan	(1056.666,1103.696867)	(1046.734488,1113.628606)
3	Feb	(1064.459501,1119.090887)	(1052.9222666,1130.627722)
4	Mar	(1065.681641,1121.552787)	(1053.882999,1133.351429)
5	Apr	(1050.171931,1095.011649)	(1040.702861,1104.480719)
6	May	(1052.690218,1092.85148)	(1044.209126,1101.332572)
7	Jun	(1062.166298,1105.244246)	(1053.069272,1114.341272)
8	Jul	(1058.913223,1101.667861)	(1049.884471,1110.696613)
9	Aug	(1063.544097,1102.283357)	(1055.363297,1110.464157)
10	Sep	(1069.21893,1106.522502)	(1061.341313,1114.400119)
11	Oct	(1074.074679,1117.965571)	(1064.805978,1127.234272)
12	Nov	(1076.462793,1128.319689)	(1065.511864,1139.270618)
13	Dec	(1065.276829,1122.762825)	(1052.722415,1135.317239)

Confidence interval is an observed interval, in principle different from sample to sample, that frequently includes the value of an unobservable parameter of interest if the experiment is repeated. If exchange rate does not fall into

these intervals, there is high probability that the exchange rate is an outlier. For example, on March, 2009, the exchange rate for 1 USD was 1452 KRW. According to the confidence interval, 1452 KRW clearly does not fall into both 95% and 99% confidence interval; therefore it was an outlier and recorded in the boxplot.

According to all the data and graphs I collected, when one needs to transfer money from South Korea to USA, it would be the best for him or her to transfer money on April or May and check if the current exchange rate either falls into confidence intervals or not.