Denco Case

Functions in R

Step1

- Data Read
  - Online Sheet → Download denco sheet as csv into projectfolder-data
  - `sales1 = read.csv("./data/denco.csv")`
  - usable copy: `sales= sales1`
- Summary/ Characteristics
  - `summary(sales)`
  - `str(sales)`
  - `unique(sales$custname)`
- Libraries
  - `library(dplyr)`
  - `library(data.table)`
  - `library(sqldf)`

Top Revenue Customers

How do I retain these customers & target incentives

M1:aggregate

```
df1 = aggregate(sales$revenue , by=list(sales$custname), FUN=sum)
```
→ `str(df1)`

```
df1[order(df1$x, decreasing=TRUE),]
```

```
head(df1[order(df1$x, decreasing=TRUE),],10)
```

	Group.1	x
1166	TRIUMPH INSULATION	3592531
268	CORNING SHARED SERVICES	12843519
1126	THERMAL PRODUCTS INC	7209418
1250	WESTERN INDUSTRIAL CERAMI	5589785
302	DELTA REFRACTORIES	5546115

M2:aggregate

Formula

```
df2 = aggregate(revenue ~ custname, data=sales, FUN=sum)
```

```
head(df2[order(df2$revenue, decreasing=T),],10)
```

	custname	revenue
1166	TRIUMPH INSULATION	3592531
268	CORNING SHARED SERVICES	12843519
1126	THERMAL PRODUCTS INC	7209418
1250	WESTERN INDUSTRIAL CERAMI	5589785
302	DELTA REFRACTORIES	5546115

M3:tapply

```
list1= tapply(sales$revenue, sales$custname, FUN=sum)
```
→ `summary(list1)`

```
head(sort(list1, decreasing=T))
```

TRIUMPH INSULATION	CORNING SHARED SERVICES	THERMAL PRODUCTS INC	WESTERN INDUSTRIAL CERAMI	DELTA REFRACTORIES	INDUSTRIAL FURNACE CO
35592531	12843519	7209418	5589785	5546115	5530092

M4: dplyr

install.packages('dplyr')

```
library(dplyr)
```
→ `names(sales)` → `"custname" "region" "partnum" "revenue" "cost" "margin"`

```
sales %>% filter(revenue > 10000000)
```

```
sales %>% group_by(custname) %>% summarize(Revenue = sum(revenue)) %>% arrange(desc(Revenue))
```

A tibble: 1,268 x 2

	custname	Revenue
	<fctr>	<dbl>
1	TRIUMPH INSULATION	35592531
2	CORNING SHARED SERVICES	12843519
3	THERMAL PRODUCTS INC	7209418
4	WESTERN INDUSTRIAL CERAMI	5589785
5	DELTA REFRACTORIES	5546115

M5:data.table

```
library(data.table)
```

```
dt1 = as.data.table(sales)
```

```
dt1[, sum(revenue), by=custname]
```

	custname	V1
1:	3M COMPANY	3106603.20
2:	4-STATE SUPPLY	404716.60
3:	A G C INC.	81982.46
4:	A P S TECHNOLOGY	333.00
5:	A W CHESTERTON COMPANY	5453.90

1264:	UNIFRAX LTD.	3376.02
1265:	UNIFRAX S.R.O.	48051.89
1266:	UTSU CO. LTD JAPAN	6691.08
1267:	VATANA PHAISAL ENGNRG CO.	29376.60
1268:	VESUVIUS MEXICO SA DE CV	871222.50

M6: sqldf

```
library(sqldf)
```

```
df5 =sqldf("Select custname, sum(revenue) from sales Group By custname order by sum(revenue) desc ")
```

```
head(df5)
```

	custname	sum(revenue)
1	TRIUMPH INSULATION	35592531
2	CORNING SHARED SERVICES	12843519
3	THERMAL PRODUCTS INC	7209418
4	WESTERN INDUSTRIAL CERAMI	5589785
5	DELTA REFRACTORIES	5546115
6	INDUSTRIAL FURNACE CO	5530092

Calculate Frequency of Purchases

Improve repeated sales, Target customers with low sales Volumes ← Who are the most loyal Customers

M1:table

`names(sales)` → `head(sort(table(sales$custname), decreasing=T))`

```
CHIZ BROS INC        ARMIL / C F S INC
          253                      174
ZAR-TECH             PYROTEK INC
          165                      164
FIBRECAST INC. REFRACTORY ENGINEERS INC.
          160                      155
```

M2:xtab

`head(sort(xtabs(~ custname, sales), decreasing=T))`

```
CHIZ BROS INC        ARMIL / C F S INC
          253                      174
ZAR-TECH             PYROTEK INC
          165                      164
FIBRECAST INC. REFRACTORY ENGINEERS INC.
          160                      155
```

M3:plyr

install.packages('plyr')

```
df2a= plyr::count(sales, c('custname'))
```
→ `str(df2a); names(df2a)`

```
head(df2a[ order(-df2a$freq),])
```

	custname	freq
237	CHIZ BROS INC	253
92	ARMIL / C F S INC	174
1265	ZAR-TECH	165
902	PYROTEK INC	164
404	FIBRECAST INC.	160
932	REFRACTORY ENGINEERS INC.	155

M4:dplyr

```
library(dplyr)
```

```
sales %>% dplyr::count(custname, sort=TRUE)
```

package::function

```
sales %>% dplyr::group_by(custname) %>% dplyr::summarise(n = n()) %>% dplyr::arrange(desc(n))
```

	custname	n
	<fctr>	<int>
1	CHIZ BROS INC	253
2	ARMIL / C F S INC	174
3	ZAR-TECH	165
4	PYROTEK INC	164
5	FIBRECAST INC.	160

Summarise by Partnum

Maximise revenue from high value parts ← What part numbers bring in to significant portion of revenue

M1:aggregate

df3a= aggregate(sales$margin, by=list(sales$partnum), FUN=sum)

```
df3a= aggregate(sales$revenue, by=list(sales$partnum), FUN=sum)
```

```
df3a[order(- df3a$x), ][1:5, ]
```

	Group.1	x
226	733648000	30075436
287	735602000	16151536
261	734370000	12641200
729	764821000	6481617
46	726000062	6309276

M2:aggregate

df3b = aggregate(margin ~ partnum, data=sales, FUN=sum)

```
df3b = aggregate(revenue ~ partnum, data=sales, FUN=sum)
```

```
head(df3b)
```

```
head(df3b[order(df3b$revenue, decreasing=T),])
```

	partnum	revenue
226	733648000	30075436
287	735602000	16151536
261	734370000	12641200
729	764821000	6481617
46	726000062	6309276
227	733649000	5735424

Highest Profit

What parts have the highest profit margin

- What parts are driving profits & what parts need to build further
- # which parts have highest Profit : partno - sum(profit)
- names(sales)
- df4a = aggregate(margin ~ partnum, data=sales, FUN=sum)
- head(df4a[order(df4a$margin, decreasing = T),])
- sales %>% group_by (partnum) %>% summarise(TotalMargin= sum(margin)) %>% arrange(desc(TotalMargin)) %>% head()

Information Required

How do I retain these customers & target incentives ← Which customers contribute the most to their revenue

See online R Code for more summarisation

https://github.com/hhnoida/analytics1/blob/master/dencoCase.R