Regression Report – David

Initial check

What number of requests don't have the same product_cd for all their inquiries? (Out of 319,733):
13,649

What number and percentage of *List requests* have this property?
Number: ***13,088***
Percentage: ***4.116 %***

What number and percentage of *individual List inquiries* have this property?
Number: ***217,503***
Percentage***: 5.078369 %***

Regressions run

```
subset <- df.inquiry %>% filter(pt == 0,
                        product_cd == "USHY",
                        p_type != "Broker-Dealer",
                        list_length >= 20,
                        numsublists < list_length/2,
                        5 < sublist_length)

regr1 <- felm(filled ~ trans_cost + mediancost_insublist + mediancost_outsidesublist |
                            req_id | 0 | req_id + date, data = subset)

regr2 <- felm(filled ~ trans_cost + mincost_insublist + mincost_outsidesublist | req_id | 0 |
                            req_id + date, data = subset)
```

Regression 1 and 2

	Estimate	Cluster.s.e.	t.value	pval
trans_cost	−1.4912	0.7656	−1.9478	0.0515
mediancost_insublist	15.8093	1.8029	8.7690	2.24e−18
mediancost_outsidesublist	31.3077	3.5764	8.7540	2.55e−18

	Estimate	Cluster.s.e.	t.value	pval
trans_cost	−7.1286	0.6272	−11.3657	0.0000
mincost_insublist	−7.1498	0.6279	−11.3874	0.0000
mincost_outsidesublist	−7.1438	0.6288	−11.3616	0.0000

RegressionSummary

David

2024-03-18

```
subset <- read_csv("~/Desktop/Portfolio_Trades_my_computer/data_minimizing/working_files/subset.csv")
```

```
## Rows: 297381 Columns: 16
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (3): product_cd, p_type, request_type
## dbl (13): req_id, req_quantity, trade_quantity, spread, filled, trans_cost, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
regr1 <- felm(filled ~ trans_cost + mediancost_insublist + mediancost_outsidesublist | req_id | 0 | req
summary1 <- summary(regr1, cluster = c("req_id"))
print(summary1)
```

```
##
## Call:
##    felm(formula = filled ~ trans_cost + mediancost_insublist + mediancost_outsidesublist |      req_
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.9980  -0.2166   0.0152   0.1990   2.9095
##
## Coefficients:
##                            Estimate Cluster s.e. t value Pr(>|t|)
## trans_cost                  -1.4912       0.7656  -1.948   0.0515 .
## mediancost_insublist        15.8093       1.8029   8.769   <2e-16 ***
## mediancost_outsidesublist   31.3077       3.5764   8.754   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3739 on 188840 degrees of freedom
##   (101568 observations deleted due to missingness)
## Multiple R-squared(full model): 0.4549   Adjusted R-squared: 0.4347
## Multiple R-squared(proj model): 0.03274   Adjusted R-squared: -0.002973
## F-statistic(full model, *iid*): 22.6 on 6972 and 188840 DF, p-value: < 2.2e-16
## F-statistic(proj model): 44.82 on 3 and 6969 DF, p-value: < 2.2e-16
```

```
regr2 <- felm(filled ~ trans_cost + mincost_insublist + mincost_outsidesublist | req_id | 0 | req_id, da
summary2 <- summary(regr2, cluster = c("req_id"))
print(summary2)
```

```
##
## Call:
##    felm(formula = filled ~ trans_cost + mincost_insublist + mincost_outsidesublist |      req_id | 0
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0701 -0.2126  0.0152  0.1941 10.5760
##
## Coefficients:
##                         Estimate Cluster s.e. t value Pr(>|t|)
## trans_cost               -7.1286       0.6272  -11.37   <2e-16 ***
## mincost_insublist        -7.1498       0.6279  -11.39   <2e-16 ***
## mincost_outsidesublist   -7.1438       0.6288  -11.36   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3647 on 188840 degrees of freedom
##   (101568 observations deleted due to missingness)
## Multiple R-squared(full model): 0.4816   Adjusted R-squared: 0.4624
## Multiple R-squared(proj model): 0.08012   Adjusted R-squared: 0.04616
## F-statistic(full model, *iid*):25.16 on 6972 and 188840 DF, p-value: < 2.2e-16
## F-statistic(proj model): 44.23 on 3 and 6969 DF, p-value: < 2.2e-16
```