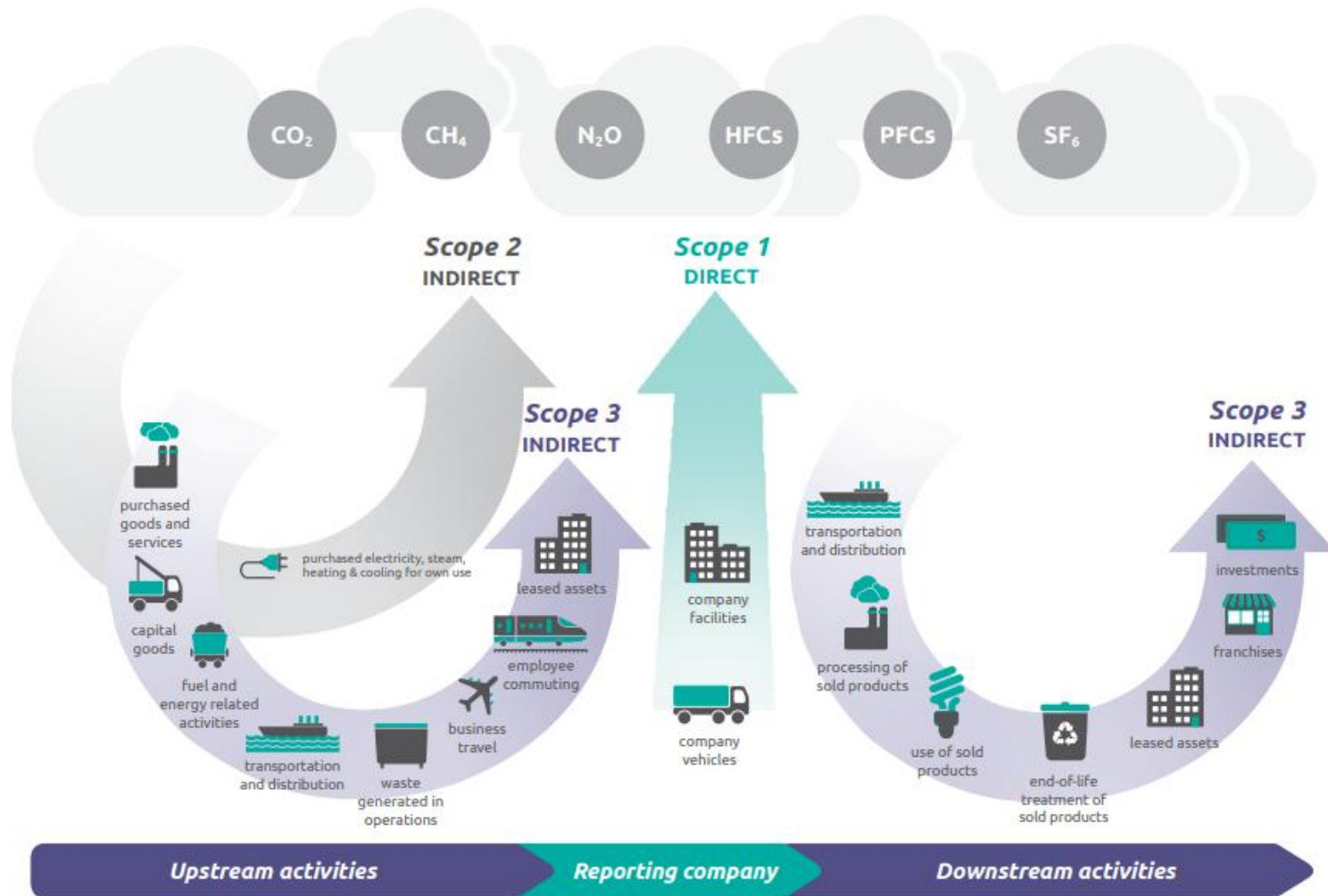# Final Manufacturer Attributes Affecting Extended Supply Chain Transparency

Abraham Kudiabor, Chulhwan Kwon,
Farshid Jahanshaheenezhad, Sagar Sinha, Uschuas Dipta Das

ASU W. P. Carey
School of Business
Arizona State University
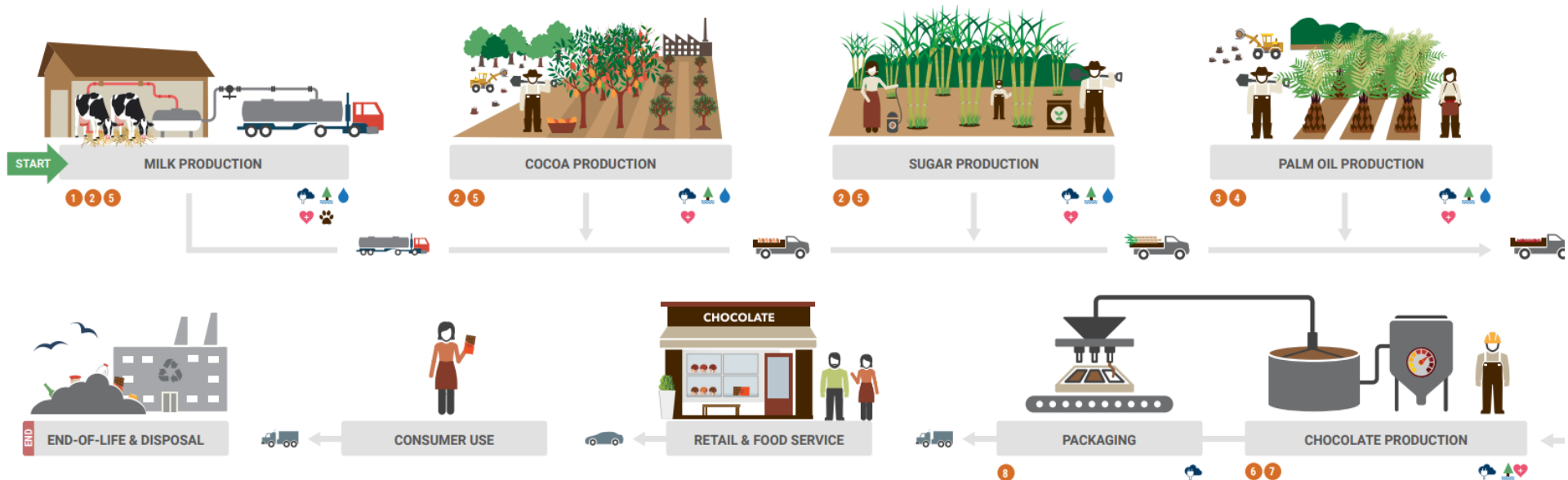
Different Scopes of emission

Increasing interest for Scope 3

- SB 253
  - California Climate Corporate Data Accountability Act
  - Requires public and some private companies to disclose Scope 1, Scope 2, and Scope 3 emissions starting 2026

- European Sustainability Reporting Standards
  - Requires public-interest entities operating in the European Union (EU), and non-EU companies with significant EU operations disclose Scope 1, 2, and 3 emissions

Lack of extended supply chain transparency

The Question

What are a buying firm's attributes that make it more likely that their upstream growers will measure and disclose Scope 1 and 2 GHG emissions?

Extended supply chain transparency

- The aggregate level of GHG emissions information that a final manufacturer's growers make available to the them and their downstream retailer.

- *Proposition: A food manufacturer will have greater extended supply chain transparency about GHG emissions when they help their growers be (1) aware that they (the buyer) want GHG emissions information, (2) motivated to be transparent, and (3) capable of measuring and disclosing GHG emissions.*

AMC model in the context of sustainability transparency

- Awareness
  - suppliers knowing that downstream buyers value and require detailed GHG emissions information

- Motivation
  - internal and external drivers that compel suppliers to engage in transparency

- Capability
  - the resources, tools, and skills necessary for effective data handling and disclosure

Source

THESIS, created and developed by TSC, is a performance assessment system that guides retailers and suppliers to benchmark, quantify, and take action on critical sustainability issues within their consumer product supply chains.

Science-Based Insights

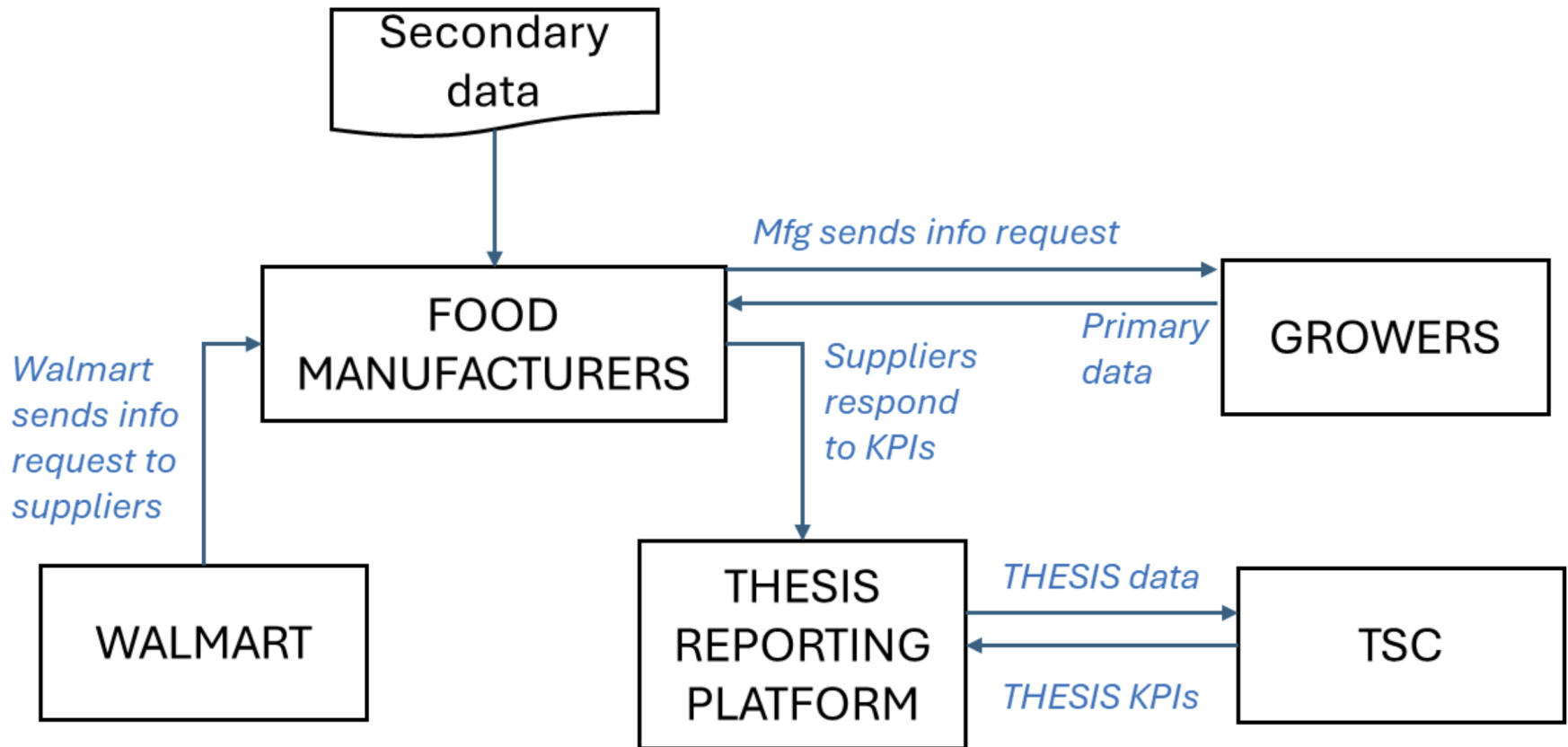Eliminates Survey Fatigue

Established Assessments

Data-Backed Decision Making

Source

- About THESIS
  - Based in Life Cycle Assessment, THESIS is a multi-stakeholder process, including business, government, NGOs, and academia.
  - THESIS reveals supply chain hotspots, which are improvement opportunities, based on a tailored survey for each product category.
  - Since 2009, TSC has been constantly updating the survey to better capture the ever-changing business landscape.

- Participants
  - 9 retailers, 1187 suppliers (51%)

Data flow in THESIS

Source

- Data used
  - 2021, 2022 food and agriculture suppliers.
    - Animal Base Foods, Apples and Pears, Beans Lentils and Peas, Beef, Berries and Grapes, Chicken, Citrus, Coffee, Complex Foods and Beverages, Cookies and Baked Goods, Cucumbers Melons and Squash, Dairy, Eggs, Farmed Fish, …, Wild-caught Shellfish

Examples

9. GREENHOUSE GAS EMISSIONS INTENSITY - ANIMAL FARM OPERATIONS

Question

What was the greenhouse gas emissions intensity associated with the animal farm operations and feed producers in your supply chain?

Response Options

A. We are unable to determine at this time.

B. Our greenhouse gas emissions intensity was:

   B1._____ kg CO_2e per metric tonne of milk.

   B2._____% of our milk supply, by mass, is represented by the number reported above.

Examples

- Calculate B1 as the average of the greenhouse gas emission intensity estimates for the animal farm operations that produced your milk supply, weighted by the mass of milk supplied by each farm.

- Calculate B2 as the mass of milk supply, for which you were able to obtain primary greenhouse gas intensity data, divided by the total mass of your milk supply then multiply by 100. If you have reported a regional estimate for B1, then report 0% for B2.

Variables

Extended Supply Chain Transparency

- Binary
- Whether the company had extended supply chain transparency

Company Size

- 1: Under 1 million USD
- 2: 1-10 million USD
- 3: 10-100 million USD
- 4: 100 million-1 billion USD
- 5: Greater than 1 billion USD

Variables

Company Type

- 0: Private Company
- 1: Public Company

Disclosure

- 0: No Disclosure
- 1: Public Disclosure

Sustainability Team Size

- 0: None or Non-Disclosure
- 1: 1 to 2
- 2: 3 to 5
- 3: 6 to 10
- 4: 11 to 20
- 5: More than 20

Information Source

- 0: Non responders, Other, We did not collect any supply chain data for the purpose of answering KPIs
- 1: We used secondary data (e.g., regional averages), We used modeled data (e.g., LCI databases)
- 2: We used previously assembled supplier data, We engaged our largest suppliers
- 3: We attempted to contact most or all of our suppliers

Past Experience

- Number of years that the company has reported THESIS

Other Experience

- Number of different environmental sustainability farm level programs the reporting company used

Control Variables

Number of Products Supplying

- The number of products the reporting company supplied to the retailers

KPI Category

- KPI categories.
- From 1 to 37

Company's Sustainability Score

- Average of all KPIs related to extended supply chain environmental transparency scores

Variables

GHG Goal Maturity

- 0: The company doesn't has a SBTi certified goal
- 1: The company has a SBTi certified goal

Contract Manufacturing

- 0: The company used a contract facility
- 1: The company only used a facility that they own

Supply Chain Transparency Interest

- 0: Others
- 1: Companies that reflected their interest in scope 3 emission or supply chain transparency

Variables

Competition

- The number of suppliers responding to the same KPI category

Number of Sourcing Continents

- The number of continents the company sourced from

Data Description

```
> psych::describe(m.data)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
KPIset_ID	1	862	1206.22	255.76	1134	1179.12	171.98	1005	3415	2410	4.61	36.12	8.71
Supplier ID	2	862	1093.11	674.73	1003	1053.15	797.64	102	2468	2366	0.40	-1.03	22.98
DV	3	862	0.28	0.45	0	0.23	0.00	0	1	1	0.97	-1.06	0.02
IV_CompSize*	4	698	2.96	0.89	3	3.02	1.48	1	4	3	-0.43	-0.68	0.03
IV_CompType	5	862	0.24	0.42	0	0.17	0.00	0	1	1	1.24	-0.45	0.01
IV_Disclosure	6	862	0.29	0.45	0	0.23	0.00	0	1	1	0.94	-1.12	0.02
IV_SusTeam*	7	862	2.48	1.35	2	2.27	1.48	1	6	5	1.15	0.83	0.05
IV_InfoSource*	8	862	2.61	1.18	3	2.64	1.48	1	4	3	-0.37	-1.40	0.04
IV_PastEx	9	862	4.15	1.28	5	4.44	0.00	0	5	5	-1.79	2.78	0.04
IV_OtherEx	10	862	0.37	0.59	0	0.28	0.00	0	4	4	1.65	3.63	0.02
C_NumbProduct	11	861	7.32	7.94	4	5.66	4.45	1	32	31	1.66	1.91	0.27
C_KPI*	12	862	17.50	10.86	15	17.25	13.34	1	37	36	0.24	-1.39	0.37
C_CorporateSus	13	797	48.78	43.00	50	48.48	74.13	0	100	100	0.04	-1.72	1.52
C_GHGMaturity	14	862	0.23	0.59	0	0.05	0.00	0	2	2	2.43	4.27	0.02
C_ContractMF	15	862	0.42	0.49	0	0.40	0.00	0	1	1	0.31	-1.90	0.02
C_SusPriorities	16	862	0.12	0.32	0	0.02	0.00	0	1	1	2.36	3.57	0.01
C_Competition	17	862	40.74	26.24	36	38.54	19.27	5	92	87	0.72	-0.55	0.89
C_NumbSourcingCont	18	665	1.11	1.56	0	0.80	0.00	0	7	7	1.54	1.80	0.06

Description

```
> Hmisc::describe((m.data[c(3, 4, 5, 6, 7, 8, 9, 10)]))
(m.data[c(3, 4, 5, 6, 7, 8, 9, 10)])

 8  Variables      862  Observations
-------------------------------------------------------------
DV
      n  missing distinct      Info      Sum      Mean
    862        0        2     0.607      243    0.2819

-------------------------------------------------------------
IV_CompSize
      n  missing distinct
    698      164        4

Value           2      3      4      5
Frequency      41    166    271    220
Proportion 0.059  0.238  0.388  0.315
-------------------------------------------------------------
IV_CompType
      n  missing distinct      Info      Sum      Mean
    862        0        2      0.54      203    0.2355

-------------------------------------------------------------
IV_Disclosure
      n  missing distinct      Info      Sum      Mean
    862        0        2     0.615      248    0.2877
```

```
-------------------------------------------------------------
IV_SusTeam
      n  missing distinct
    862        0        6

Value           0      1      2      3      4      5
Frequency     201    334    183     60     28     56
Proportion  0.233  0.387  0.212  0.070  0.032  0.065
-------------------------------------------------------------
IV_InfoSource
      n  missing distinct
    862        0        4

Value           0      1      2      3
Frequency     275     11    349    227
Proportion  0.319  0.013  0.405  0.263
-------------------------------------------------------------
IV_PastEx
      n  missing distinct      Info      Mean   pMedian       Gmd
    862        0        6     0.808     4.151       4.5     1.199

Value           0      1      2      3      4      5
Frequency      38      9     42     93    194    486
Proportion  0.044  0.010  0.049  0.108  0.225  0.564

For the frequency table, variable is rounded to the nearest 0
-------------------------------------------------------------
IV_OtherEx
      n  missing distinct      Info      Mean   pMedian       Gmd
    862        0        5     0.666    0.3689       0.5    0.5284

Value           0      1      2      3      4
Frequency     585    241     33      1      2
Proportion  0.679  0.280  0.038  0.001  0.002
```
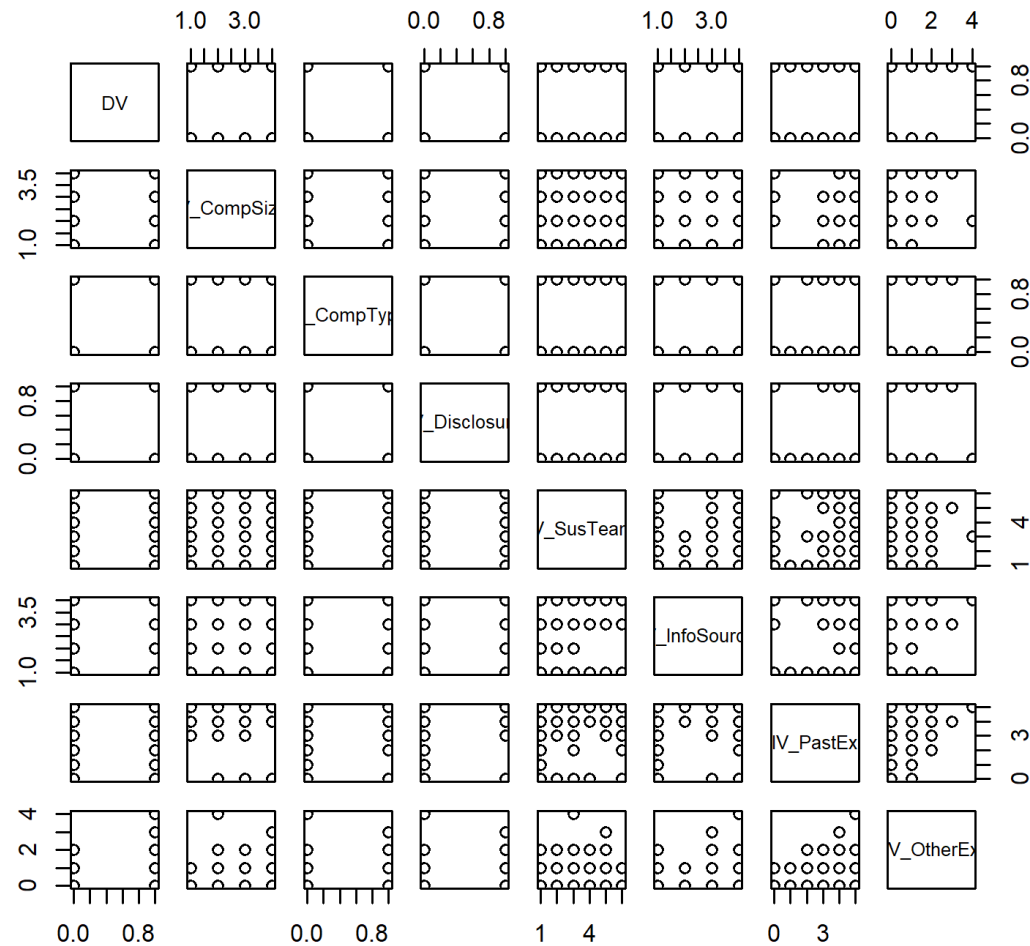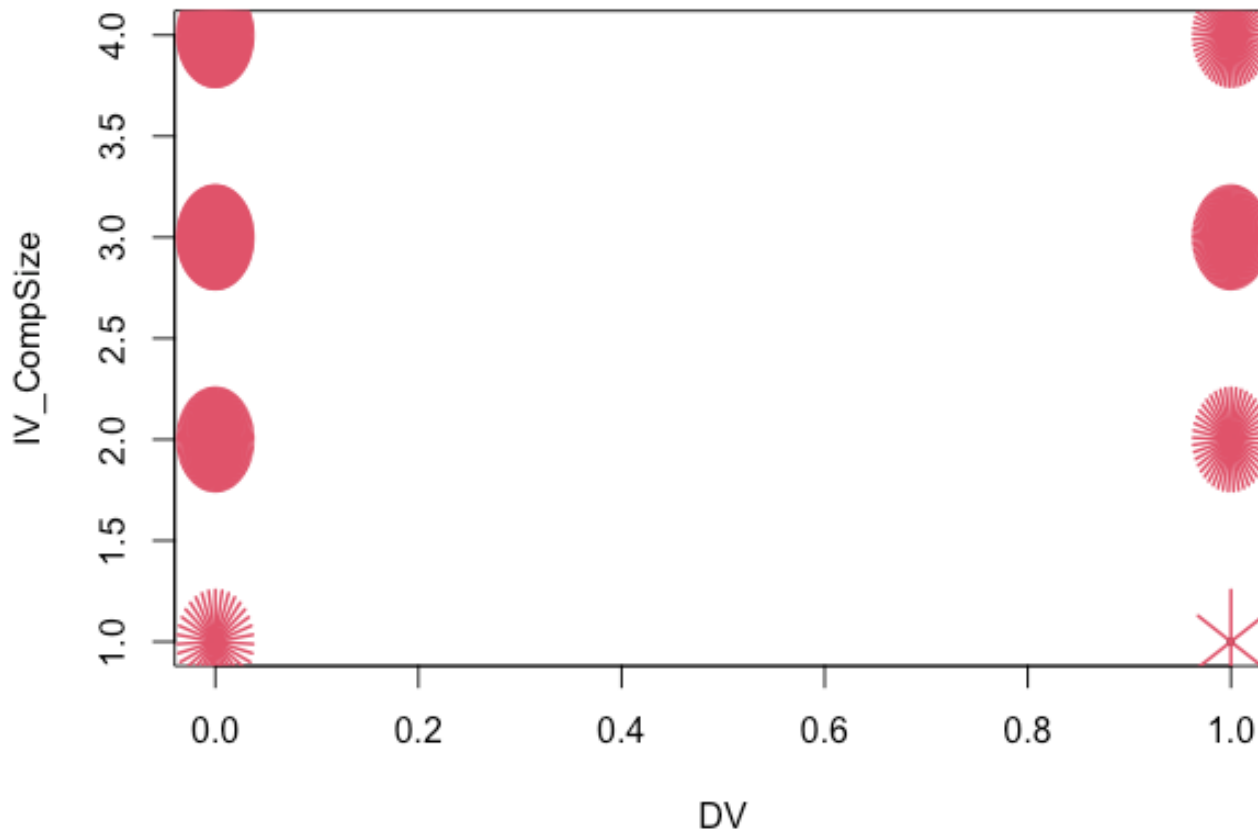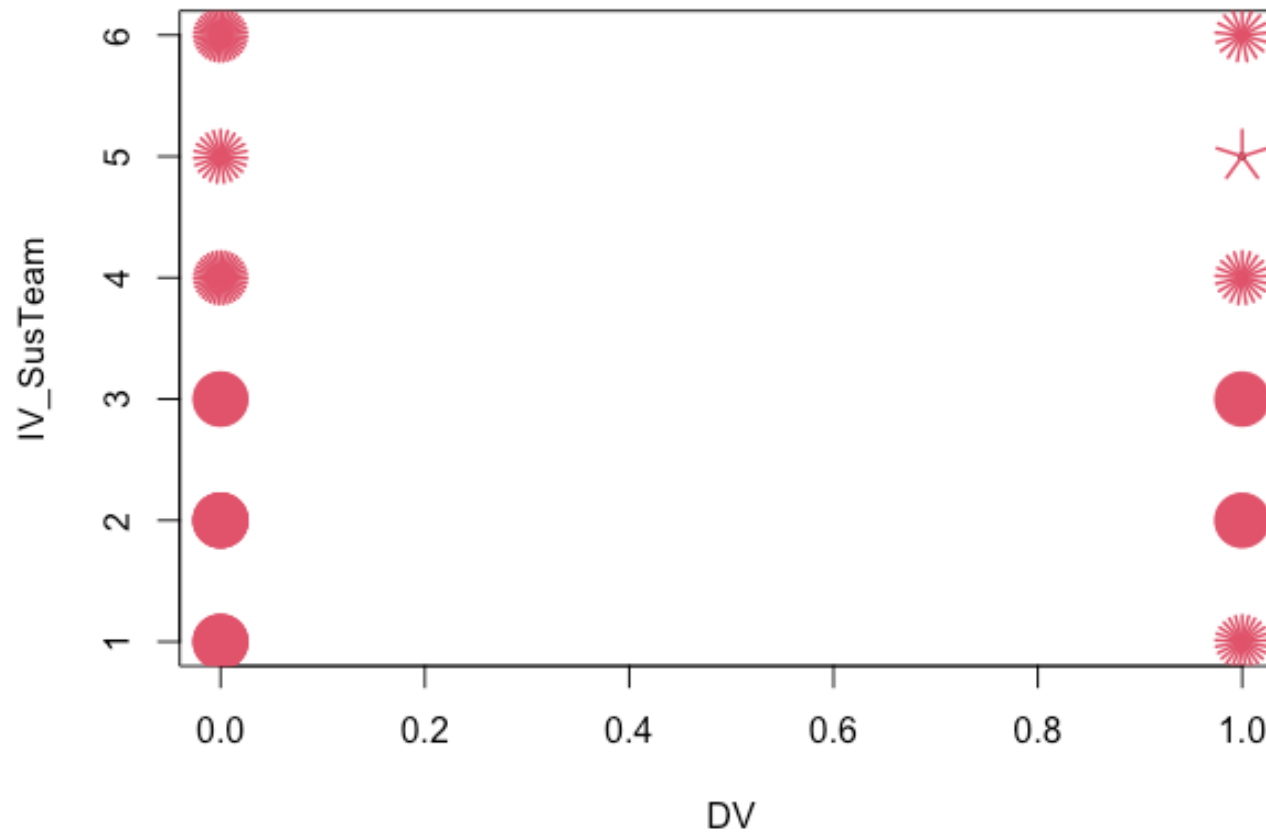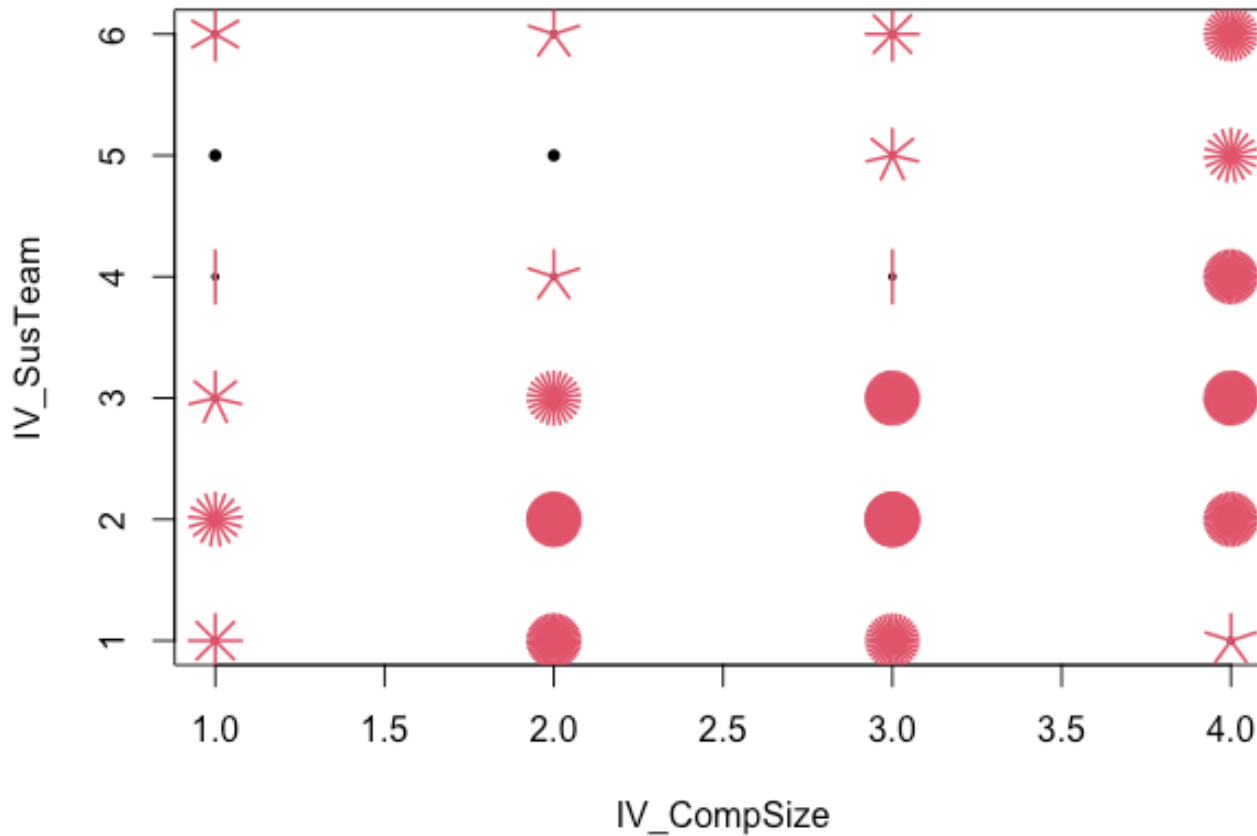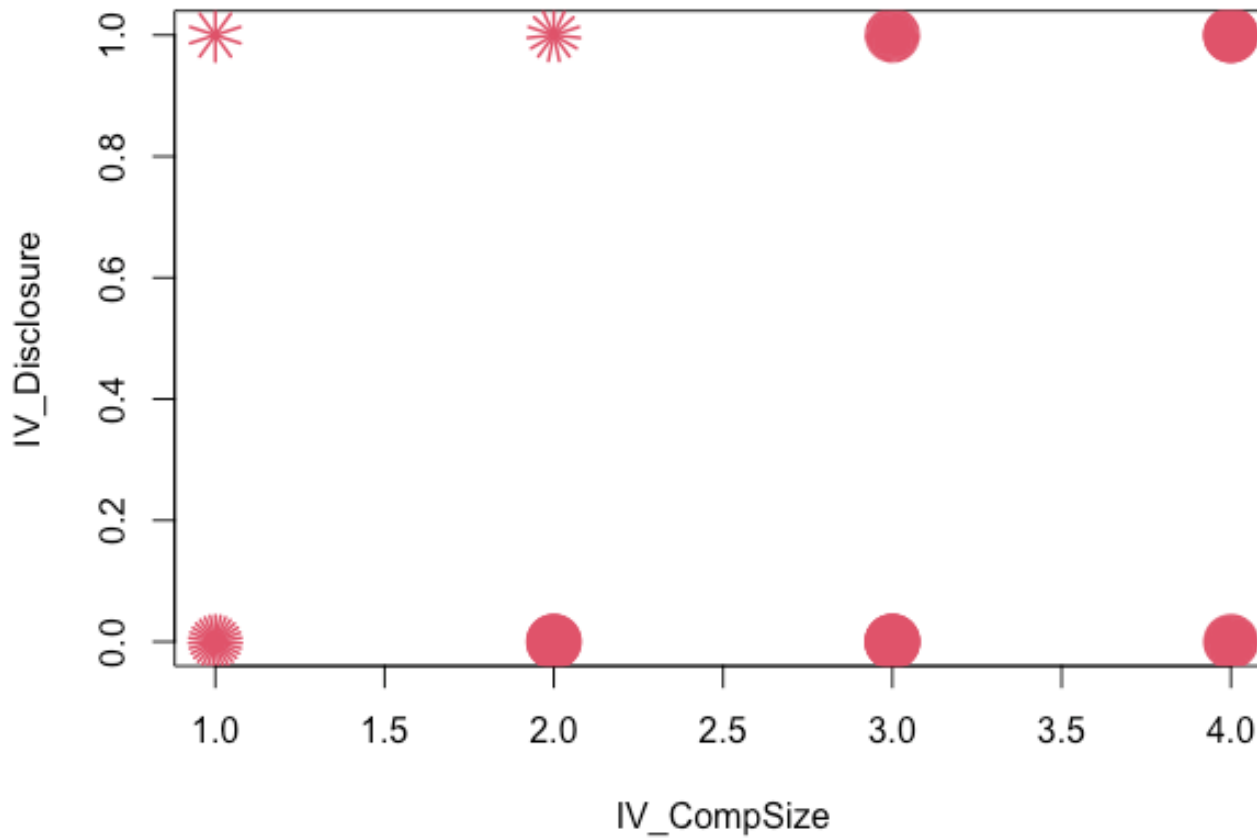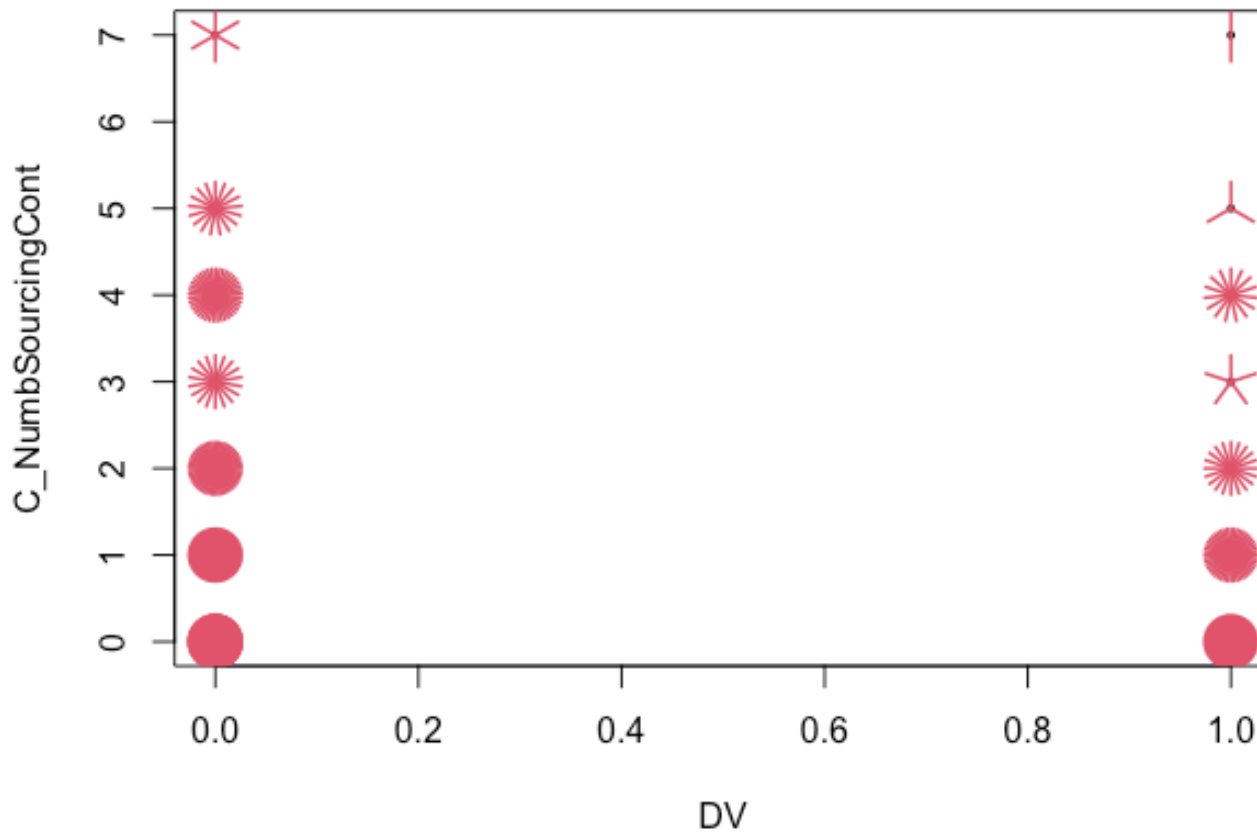
Description

Description

Description

Description

Description

Description

Data

Correlation Matrix

	DV	IV_CompType	IV_Disclosure	IV_PastEx	IV_OtherEx	C_GHGMaturity	C_ContractMF	C_SusPriorities	C_Competition
DV	1.000	-0.007	0.006	0.105	0.252	0.041	0.138	0.002	0.038
IV_CompType	-0.007	1.000	0.541	0.073	-0.102	0.432	0.212	0.084	0.020
IV_Disclosure	0.006	0.541	1.000	0.227	-0.055	0.409	0.074	0.156	-0.058
IV_PastEx	0.105	0.073	0.227	1.000	-0.023	0.190	0.004	0.119	-0.089
IV_OtherEx	0.252	-0.102	-0.055	-0.023	1.000	-0.193	0.187	-0.010	0.191
C_GHGMaturity	0.041	0.432	0.409	0.190	-0.193	1.000	-0.039	0.031	-0.112
C_ContractMF	0.138	0.212	0.074	0.004	0.187	-0.039	1.000	0.109	0.082
C_SusPriorities	0.002	0.084	0.156	0.119	-0.010	0.031	0.109	1.000	0.149
C_Competition	0.038	0.020	-0.058	-0.089	0.191	-0.112	0.082	0.149	1.000

Listwise deletion

```
vars_for_model <- c("DV", "IV_CompSize", "IV_CompType", "IV_Disclosure",
                    "IV_SusTeam", "IV_InfoSource", "IV_PastEx",
                    "IV_OtherEx", "C_NumbProduct", "C_KPI",
                    "C_CorporateSus", "C_GHGMaturity", "C_ContractMF",
                    "C_SusPriorities", "C_NumbSourcingCont", "C_Competition")

cleaned_data <- na.omit(m.data[ , vars_for_model ])
```

```
> str(cleaned_data)
tibble [510 × 16] (S3: tbl_df/tbl/data.frame)
 $ DV                : num [1:510] 0 0 1 0 0 0 1 0 1 1 ...
 $ IV_CompSize       : Factor w/ 4 levels "2","3","4","5": 2 4 3 4 2 1 3 2 3 3 ...
 $ IV_CompType       : num [1:510] 0 0 0 1 0 0 0 1 0 1 ...
 $ IV_Disclosure     : num [1:510] 0 0 1 1 1 0 0 0 0 0 ...
 $ IV_SusTeam        : Factor w/ 6 levels "0","1","2","3",..: 1 2 3 3 3 2 2 2 3 2 ...
 $ IV_InfoSource     : Factor w/ 4 levels "0","1","2","3": 4 1 1 4 4 4 3 1 4 3 ...
 $ IV_PastEx         : num [1:510] 5 5 5 5 4 4 4 3 5 4 ...
 $ IV_OtherEx        : num [1:510] 0 0 0 0 0 0 0 0 1 0 ...
 $ C_NumbProduct     : num [1:510] 1 25 2 4 1 1 2 3 13 3 ...
 $ C_KPI             : Factor w/ 37 levels "1","2","3","4",..: 1 1 1 1 1 1 1 1 2 2 ...
 $ C_CorporateSus    : num [1:510] 17 0 33
 $ C_GHGMaturity     : num [1:510] 0 0 0 0
 $ C_ContractMF      : num [1:510] 0 1 1 1
 $ C_SusPriorities   : num [1:510] 1 1 1 1
 $ C_NumbSourcingCont: num [1:510] 0 2 7 0
 $ C_Competition     : num [1:510] 24 24 24
 - attr(*, "na.action")= 'omit' Named int [
  ..- attr(*, "names")= chr [1:352] "3" "4"
```

- Missing data were missing completely at random.
- Initially, 862 samples, reduced to 509 after listwise deletion.

Logistic regression only with independent variables

```
Call:
glm(formula = DV ~ IV_CompSize + IV_CompType + IV_Disclosure +
    IV_SusTeam + IV_InfoSource + IV_PastEx + IV_OtherEx, family = binomial,
    data = cleaned_data)

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)       -6.8708     1.0559  -6.507 7.65e-11 ***
IV_CompSize3       1.9190     0.8180   2.346 0.018975 *
IV_CompSize4       1.7263     0.7974   2.165 0.030406 *
IV_CompSize5       1.4379     0.8449   1.702 0.088774 .
IV_CompType        0.3381     0.3301   1.024 0.305739
IV_Disclosure     -0.4728     0.3260  -1.450 0.146930
IV_SusTeam1        0.9716     0.5216   1.863 0.062512 .
IV_SusTeam2        3.0075     0.5429   5.539 3.03e-08 ***
IV_SusTeam3        2.3247     0.6686   3.477 0.000507 ***
IV_SusTeam4        1.2755     0.7770   1.642 0.100690
IV_SusTeam5        2.5358     0.6429   3.944 8.01e-05 ***
IV_InfoSource1   -12.8068   515.6921  -0.025 0.980187
IV_InfoSource2     0.5819     0.3717   1.565 0.117479
IV_InfoSource3     0.6030     0.3662   1.646 0.099678 .
IV_PastEx          0.4082     0.1241   3.290 0.001002 **
IV_OtherEx         0.7879     0.1929   4.086 4.40e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 617.91  on 509  degrees of freedom
Residual deviance: 478.29  on 494  degrees of freedom
AIC: 510.29
```

```
> vif(nm1)
                 GVIF Df GVIF^(1/(2*Df))
IV_CompSize   3.205432  3        1.214268
IV_CompType   1.966530  1        1.402330
IV_Disclosure 2.043767  1        1.429604
IV_SusTeam    2.309565  5        1.087309
IV_InfoSource 1.640364  3        1.085984
IV_PastEx     1.331234  1        1.153791
IV_OtherEx    1.258466  1        1.121814
```

```
> # Likelihood-Ratio Test of global model utility
> pchisq((617.91 - 478.29), df=(509 - 494), lower.tail=F)
[1] 2.735262e-22
```

Logistic regression with independent and control variables (1)

```
> nm2 <- glm(DV ~ IV_CompSize + IV_CompType + IV_Disclosure  + IV_SusTeam +  IV_InfoSource +
+            IV_PastEx + IV_OtherEx + C_NumbProduct + C_KPI + C_CorporateSus + C_GHGMaturity +
+            C_ContractMF + C_SusPriorities + C_Competition + C_NumbSourcingCont,
+            data=cleaned_data, family=binomial)
경고메시지(들):
glm.fit: 적합된 확률값들이 0 또는 1 입니다
```

Warning indicating perfect separation or quasi-complete separation:

glm.fit: fitted probabilities numerically 0 or 1 occurred.

Logistic regression with independent and control variables (1)

```
> # Cross-tabulations to check for separation with categorical predictors
> table(cleaned_data$C_KPI, cleaned_data$DV)

      0  1
1     6  2        21  8  0
2     7 11        22 12 16
3     5  2        23  0  0
4    16  1        24  2  0
5     3  0        25  0  0
6    43 21        26 25  6
7    15  0        27  7  1
8    21  4        28  9  7
9     3  1        29  3  0
10   41  0        30 27  0
11    9  3        31 17 11
12    9 12        32  0  0
13   15  7        33 10 22
14    3  0        34 15 14
15    1  4        35  4  1
16    0  0        36  0  0
17   16  1        37  0  0
18    8  3
19    0  0
20    0  0
```

Logistic regression excluding "KPI" (2)

```
glm(formula = DV ~ IV_CompSize + IV_CompType + IV_Disclosure +
    IV_SusTeam + IV_InfoSource + IV_PastEx + IV_OtherEx + C_NumbProduct +
    C_CorporateSus + C_GHGMaturity + C_ContractMF + C_SusPriorities +
    C_Competition + C_NumbSourcingCont, family = binomial, data = cleaned_data)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)      -6.663538   1.200589  -5.550 2.85e-08 ***
IV_CompSize3      1.119967   0.841738   1.331  0.18334
IV_CompSize4      0.964920   0.842833   1.145  0.25227
IV_CompSize5      0.431061   0.944033   0.457  0.64795
IV_CompType       0.087857   0.416134   0.211  0.83279
IV_Disclosure    -0.970506   0.393248  -2.468  0.01359 *
IV_SusTeam1       0.921485   0.555137   1.660  0.09693 .
IV_SusTeam2       2.808094   0.601220   4.671 3.00e-06 ***
IV_SusTeam3       1.418484   0.854289   1.660  0.09683 .
IV_SusTeam4       1.238494   0.907487   1.365  0.17233
IV_SusTeam5       2.066768   0.746458   2.769  0.00563 **
IV_InfoSource1  -14.224920 789.987623  -0.018  0.98563
IV_InfoSource2   -0.001560   0.446831  -0.003  0.99722
IV_InfoSource3    0.139489   0.440635   0.317  0.75157
IV_PastEx         0.343279   0.138996   2.470  0.01352 *
IV_OtherEx        0.697751   0.228980   3.047  0.00231 **
C_NumbProduct     0.050091   0.022918   2.186  0.02884 *
C_CorporateSus    0.024938   0.003659   6.816 9.36e-12 ***
C_GHGMaturity     0.502023   0.325801   1.541  0.12334
C_ContractMF      0.453304   0.327338   1.385  0.16611
C_SusPriorities   0.095413   0.360550   0.265  0.79129
C_Competition    -0.007516   0.005956  -1.262  0.20694
C_NumbSourcingCont -0.146299  0.136975  -1.068  0.28549
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 617.91  on 509  degrees of freedom
Residual deviance: 406.26  on 487  degrees of freedom
```

- Disclosure: 0.38
- Sustainability Team 2: 16.58
- Sustainability Team 5: 7.90
- Past Experience: 1.41
- Other Experience: 2.01

- Sustainability Team 1: 2.51
- Sustainability Team 3: 4.13

Logistic regression excluding "KPI" (2)

```
> # Likelihood-Ratio Test of global model utility
> pchisq((617.91 - 406.26), df=(509 - 487), lower.tail=F)
[1] 5.882385e-33


> vif(nm3)
                      GVIF Df GVIF^(1/(2*Df))
IV_CompSize       4.685532  3        1.293580
IV_CompType       2.567684  1        1.602400
IV_Disclosure     2.448951  1        1.564912
IV_SusTeam        8.539122  5        1.239200
IV_InfoSource     2.067131  3        1.128655
IV_PastEx         1.434819  1        1.197839
IV_OtherEx        1.486942  1        1.219402
C_NumbProduct     2.930404  1        1.711842
C_CorporateSus    1.345627  1        1.160012
C_GHGMaturity     3.068571  1        1.751734
C_ContractMF      1.625095  1        1.274792
C_SusPriorities   1.238087  1        1.112694
C_Competition     1.446894  1        1.202869
C_NumbSourcingCont 2.783793 1        1.668470
```

Hosmer–Lemeshow test

```
> hoslem.test(cleaned_data$DV, preds1, g = 10)

        Hosmer and Lemeshow goodness of fit (GOF) test

data:  cleaned_data$DV, preds1
X-squared = 19.226, df = 8, p-value = 0.0137


>
> preds2 <- predict(nm3, type = "response")
> hoslem.test(cleaned_data$DV, preds2, g = 10)

        Hosmer and Lemeshow goodness of fit (GOF) test

data:  cleaned_data$DV, preds2
X-squared = 3.4583, df = 8, p-value = 0.9024
```
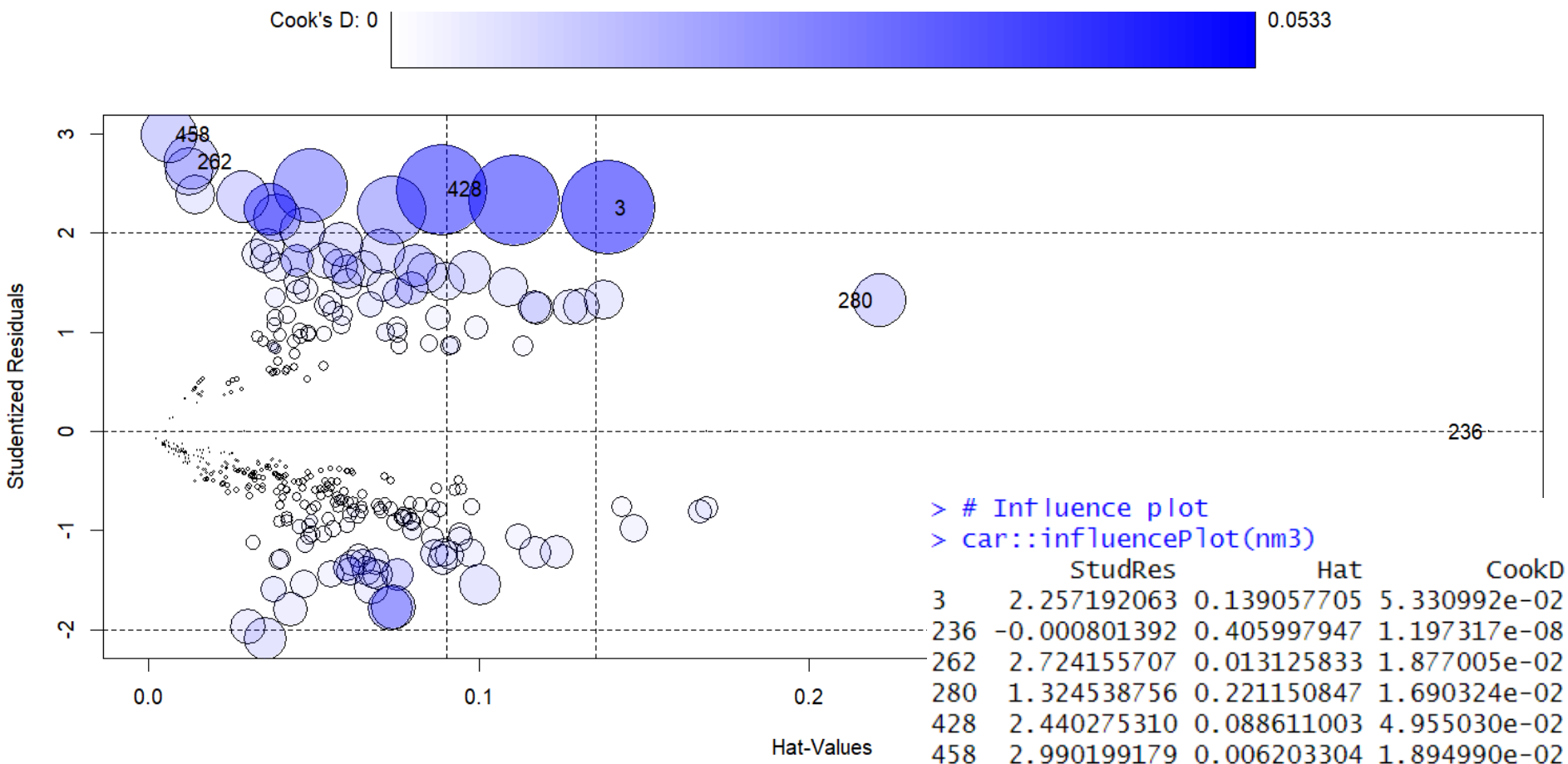
Diagnostics

Pseudo R-squares

- McFadden's Pseudo R-square

```
> 1-nm3$deviance/nm3$null.deviance
[1] 0.3425342
```

- Tjur'a Pseudo R-square

```
> abs(mean(predict(nm3, type="response")[sel]) -
+       mean(predict(nm3, type="response")[!sel]))
[1] 0.3860639
```

Influence Plot


```
> # Influence plot
> car::influencePlot(nm3)
          StudRes          Hat         CookD
3      2.257192063  0.139057705  5.330992e-02
236   -0.000801392  0.405997947  1.197317e-08
262    2.724155707  0.013125833  1.877005e-02
280    1.324538756  0.221150847  1.690324e-02
428    2.440275310  0.088611003  4.955030e-02
458    2.990199179  0.006203304  1.894990e-02
```

Cook's D

```
> max(pf(q = cooks.distance(nm3), df1 = p, df2 = n - p)*100)
[1] 1.864092e-09
> round(max(pf(q = cooks.distance(nm3), df1 = p, df2 = n - p)*100), 5)
[1] 0
> min(pf(q = cooks.distance(nm3), df1 = p, df2 = n - p)*100)
[1] 8.314044e-134


Row 3 with Cook's distance: 0.05330992
Row 428 with Cook's distance: 0.0495503
Row 369 with Cook's distance: 0.04941998
Row 7 with Cook's distance: 0.03324705
Row 474 with Cook's distance: 0.02905321
Row 458 with Cook's distance: 0.0189499
Row 262 with Cook's distance: 0.01877005
Row 280 with Cook's distance: 0.01690324
Row 207 with Cook's distance: 0.01652432
Row 244 with Cook's distance: 0.01612859
```

Discussion

What we expected

- Having a sustainability team was found to increase the odds of having extended supply chain transparency. Holding all else constant, having 3 to 5 sustainability employees showed 16.58 higher odds, and having more than 20 employees showed 7.9 higher odds of having extended supply chain transparency. While not under the $\alpha < 0.05$ (but less than 0.1), having 1 to 2 employees and having 6 to 10 employees also showed 2.51 and 4.13 higher odds.

- Experience also was found to affect the odds of having extended supply chain transparency. Holding all else constant, 1 year increase in experience of reporting to THESIS increased the odds by 1.41, and 1 unit increase in reporting to other tools increased the odds by 2.01.

What we didn't expect

- However, while having a larger sustainability team was expected to show highest odds, having 3 to 5 employees was found to have the highest odds of having extended supply chain transparency.

- Also, while having a sustainability report was thought to increase the odds of having extended supply chain transparency, the analysis results showed the odds decreasing by 62%, holding all else equal.

Q&A