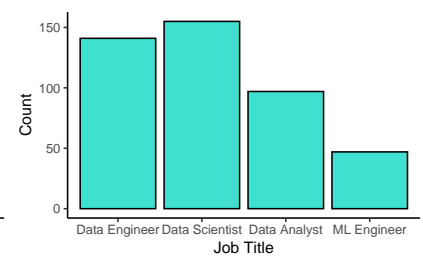
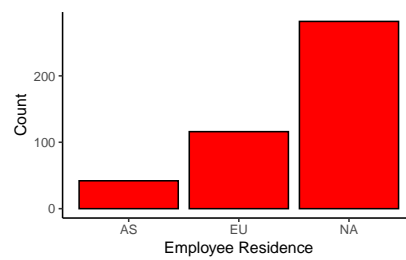
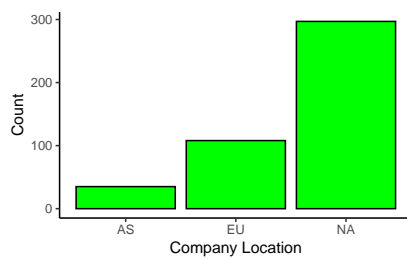
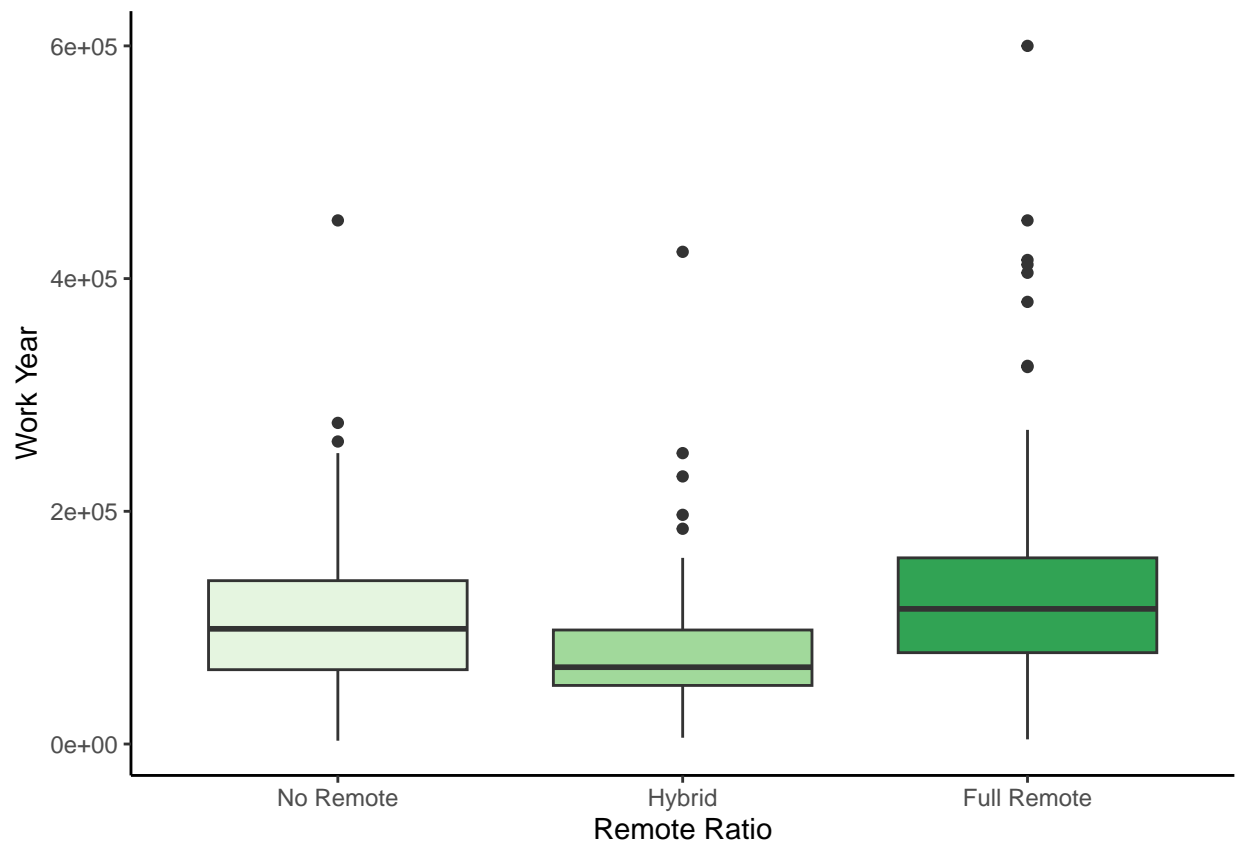


data_analysis

Wafiakmal Miftah

2022-12-03





y.level	term	estimate	std.error	statistic	p.value	conf.low	conf.high
Hybrid	(Intercept)	-0.34	0.63	-0.53	0.59	-1.58	0.90
Hybrid	experience_levelEX	0.27	0.88	0.30	0.76	-1.46	2.00
Hybrid	experience_levelMI	-0.54	0.45	-1.20	0.23	-1.41	0.34
Hybrid	experience_levelSE	-0.42	0.48	-0.87	0.38	-1.36	0.52
Hybrid	employment_typeFL	1.90	1.08	1.76	0.08	-0.21	4.02
Hybrid	employment_typeFT	1.06	0.54	1.95	0.05	-0.01	2.13
Hybrid	employment_typePT	16.98	0.37	46.44	0.00	16.27	17.70
Hybrid	job_titleData Scientist	0.36	0.39	0.93	0.35	-0.40	1.12
Hybrid	job_titleData Analyst	0.17	0.49	0.34	0.73	-0.79	1.13
Hybrid	job_titleML Engineer	1.17	0.52	2.23	0.03	0.14	2.20
Hybrid	employee_residenceEU	-1.33	1.31	-1.01	0.31	-3.90	1.25
Hybrid	employee_residenceNA	-2.19	1.43	-1.52	0.13	-5.00	0.62
Hybrid	company_locationEU	2.00	1.34	1.49	0.14	-0.63	4.64
Hybrid	company_locationNA	1.39	1.43	0.97	0.33	-1.42	4.20
Hybrid	company_sizeM	-2.10	0.36	-5.82	0.00	-2.81	-1.39
Hybrid	company_sizeS	-1.10	0.46	-2.41	0.02	-2.00	-0.20
Full Remote	(Intercept)	14.31	0.55	25.91	0.00	13.22	15.39
Full Remote	experience_levelEX	0.48	0.72	0.67	0.50	-0.93	1.89
Full Remote	experience_levelMI	-0.45	0.37	-1.22	0.22	-1.18	0.27
Full Remote	experience_levelSE	-0.01	0.38	-0.03	0.97	-0.76	0.74
Full Remote	employment_typeFL	-13.83	0.92	-15.05	0.00	-15.63	-12.03
Full Remote	employment_typeFT	-13.56	0.47	-28.75	0.00	-14.49	-12.64
Full Remote	employment_typePT	0.95	0.37	2.61	0.01	0.24	1.67
Full Remote	job_titleData Scientist	-0.46	0.26	-1.78	0.07	-0.97	0.05
Full Remote	job_titleData Analyst	0.07	0.30	0.22	0.82	-0.52	0.66
Full Remote	job_titleML Engineer	-0.03	0.42	-0.07	0.95	-0.85	0.79
Full Remote	employee_residenceEU	-1.77	1.26	-1.40	0.16	-4.23	0.70
Full Remote	employee_residenceNA	-2.13	1.32	-1.61	0.11	-4.72	0.46
Full Remote	company_locationEU	2.17	1.30	1.67	0.09	-0.37	4.71
Full Remote	company_locationNA	3.05	1.34	2.28	0.02	0.42	5.68
Full Remote	company_sizeM	-0.21	0.25	-0.82	0.41	-0.70	0.29
Full Remote	company_sizeS	-0.04	0.39	-0.11	0.91	-0.80	0.71

Likelihood ratio tests of Multinomial Models
##

y.level	term	estimate	std.error	statistic	p.value	conf.low	conf.high
Hybrid	(Intercept)	0.71	0.63	-0.53	0.59	0.21	2.47
Hybrid	experience_levelEX	1.31	0.88	0.30	0.76	0.23	7.37
Hybrid	experience_levelMI	0.59	0.45	-1.20	0.23	0.24	1.41
Hybrid	experience_levelSE	0.66	0.48	-0.87	0.38	0.26	1.69
Hybrid	employment_typeFL	6.71	1.08	1.76	0.08	0.81	55.63
Hybrid	employment_typeFT	2.89	0.54	1.95	0.05	0.99	8.41
Hybrid	employment_typePT	23760119.21	0.37	46.44	0.00	11603049.15	48654733.58
Hybrid	job_titleData Scientist	1.43	0.39	0.93	0.35	0.67	3.06
Hybrid	job_titleData Analyst	1.18	0.49	0.34	0.73	0.45	3.09
Hybrid	job_titleML Engineer	3.22	0.52	2.23	0.03	1.15	9.00
Hybrid	employee_residenceEU	0.27	1.31	-1.01	0.31	0.02	3.48
Hybrid	employee_residenceNA	0.11	1.43	-1.52	0.13	0.01	1.87
Hybrid	company_locationEU	7.41	1.34	1.49	0.14	0.53	103.09
Hybrid	company_locationNA	4.01	1.43	0.97	0.33	0.24	66.64
Hybrid	company_sizeM	0.12	0.36	-5.82	0.00	0.06	0.25
Hybrid	company_sizeS	0.33	0.46	-2.41	0.02	0.14	0.82
Full Remote	(Intercept)	1633491.10	0.55	25.91	0.00	553410.22	4821546.59
Full Remote	experience_levelEX	1.62	0.72	0.67	0.50	0.40	6.62
Full Remote	experience_levelMI	0.64	0.37	-1.22	0.22	0.31	1.32
Full Remote	experience_levelSE	0.99	0.38	-0.03	0.97	0.47	2.09
Full Remote	employment_typeFL	0.00	0.92	-15.05	0.00	0.00	0.00
Full Remote	employment_typeFT	0.00	0.47	-28.75	0.00	0.00	0.00
Full Remote	employment_typePT	2.60	0.37	2.61	0.01	1.27	5.32
Full Remote	job_titleData Scientist	0.63	0.26	-1.78	0.07	0.38	1.05
Full Remote	job_titleData Analyst	1.07	0.30	0.22	0.82	0.59	1.93
Full Remote	job_titleML Engineer	0.97	0.42	-0.07	0.95	0.43	2.21
Full Remote	employee_residenceEU	0.17	1.26	-1.40	0.16	0.01	2.02
Full Remote	employee_residenceNA	0.12	1.32	-1.61	0.11	0.01	1.58
Full Remote	company_locationEU	8.72	1.30	1.67	0.09	0.69	110.64
Full Remote	company_locationNA	21.19	1.34	2.28	0.02	1.53	293.85
Full Remote	company_sizeM	0.81	0.25	-0.82	0.41	0.49	1.33
Full Remote	company_sizeS	0.96	0.39	-0.11	0.91	0.45	2.04

```
## Response: remote_ratio
```

```
##
```

```
## 1          experience_level + employment_type + job_title + salary_in_usd + employee_residence
```

```
## 2 experience_level + employment_type + job_title + salary_in_usd + employee_residence + company_location
```

```
##   Resid. df Resid. Dev   Test    Df LR stat.      Pr(Chi)
```

```
## 1       850    712.4610
```

```
## 2       846    671.1218 1 vs 2     4 41.33917 2.286459e-08
```

```
## Call:
```

```
## multinom(formula = remote_ratio ~ experience_level + employment_type +
```

```
##   job_title + salary_in_usd + employee_residence + company_location,
```

```
##   data = train, trace = FALSE)
```

```
##
```

```
## Coefficients:
```

```
##          (Intercept) experience_levelEX experience_levelMI
```

```
## Hybrid          -0.9718764           0.6077138          -0.5616949
```

```
## Full Remote 13.3927538           0.1848310          -0.2951606
```

```
##          experience_levelSE employment_typeFL employment_typeFT
```

Table 1: Regression Result

	<i>Dependent variable:</i>	
	Hybrid	Full Remote
	(1)	(2)
experience_levelEX	0.27 (0.88)	0.48 (0.72)
experience_levelMI	-0.54 (0.45)	-0.45 (0.37)
experience_levelSE	-0.42 (0.48)	-0.01 (0.38)
employment_typeFL	1.90* (1.08)	-13.83*** (0.92)
employment_typeFT	1.06* (0.54)	-13.56*** (0.47)
employment_typePT	16.98*** (0.37)	0.95*** (0.37)
job_titleData Scientist	0.36 (0.39)	-0.46* (0.26)
job_titleData Analyst	0.17 (0.49)	0.07 (0.30)
job_titleML Engineer	1.17** (0.52)	-0.03 (0.42)
employee_residenceEU	-1.33 (1.31)	-1.77 (1.26)
employee_residenceNA	-2.19 (1.43)	-2.13 (1.32)
company_locationEU	2.00 (1.34)	2.17* (1.30)
company_locationNA	1.39 (1.43)	3.05** (1.34)
company_sizeM	-2.10*** (0.36)	-0.21 (0.25)
company_sizeS	-1.10** (0.46)	-0.04 (0.39)
Constant	-0.34 (0.63)	14.31*** (0.55)
Akaike Inf. Crit.	975.64	975.64
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01		

type	group	term	contrast	estimate	std.error	statistic	p.value	conf.low	conf.high
probs	No Remote	job_title	Data Scientist - Data Engineer	0.05	0.04	1.27	0.20	-0.03	0.13
probs	No Remote	job_title	Data Analyst - Data Engineer	-0.01	0.04	-0.27	0.78	-0.10	0.08
probs	No Remote	job_title	ML Engineer - Data Engineer	-0.03	0.06	-0.52	0.60	-0.14	0.08
probs	Hybrid	job_title	Data Scientist - Data Engineer	0.07	0.03	2.08	0.04	0.00	0.14
probs	Hybrid	job_title	Data Analyst - Data Engineer	0.01	0.04	0.27	0.79	-0.06	0.08
probs	Hybrid	job_title	ML Engineer - Data Engineer	0.13	0.05	2.60	0.01	0.03	0.23
probs	Full Remote	job_title	Data Scientist - Data Engineer	-0.12	0.05	-2.58	0.01	-0.21	0.09
probs	Full Remote	job_title	Data Analyst - Data Engineer	0.00	0.05	0.04	0.96	-0.10	0.10
probs	Full Remote	job_title	ML Engineer - Data Engineer	-0.10	0.07	-1.46	0.14	-0.23	0.03

Items	Value
Accuracy	0.66
Kappa	0.23
AccuracyLower	0.62
AccuracyUpper	0.70
AccuracyNull	0.63
AccuracyPValue	0.04
McnemarPValue	0.00

```
## Hybrid          -0.6607227      0.8236287      0.5484802
## Full Remote    -0.2278289     -13.1215378     -12.8737933
##               employment_typePT job_titleData Scientist job_titleData Analyst
## Hybrid          15.4915767      0.8399116      0.6943165
## Full Remote    -0.3722455     -0.1616398      0.4658540
```

```

##          job_titleML Engineer salary_in_usd employee_residenceEU
## Hybrid          1.6591688 -2.424504e-06          0.4541506
## Full Remote      -0.1101084  1.141724e-06          -1.9408026
##          employee_residenceNA company_locationEU company_locationNA
## Hybrid          -0.3147699          0.2119898          -0.1799808
## Full Remote      -2.0402946          2.0324016          2.8073034
##
## Std. Errors:
##          (Intercept) experience_levelEX experience_levelMI
## Hybrid          1.123412e-11          1.145038e-12          3.639429e-12
## Full Remote      5.313338e-12          4.649171e-13          1.396824e-12
##          experience_levelSE employment_typeFL employment_typeFT
## Hybrid          4.873124e-12          6.412531e-14          1.115786e-11
## Full Remote      3.678284e-12          2.166571e-14          5.295621e-12
##          employment_typePT job_titleData Scientist job_titleData Analyst
## Hybrid          1.219105e-14          5.319803e-12          1.451573e-12
## Full Remote      1.219102e-14          2.019548e-12          8.143862e-13
##          job_titleML Engineer salary_in_usd employee_residenceEU
## Hybrid          2.203639e-12  1.462085e-06          4.316562e-12
## Full Remote      7.841697e-13  9.022475e-07          1.377671e-12
##          employee_residenceNA company_locationEU company_locationNA
## Hybrid          6.076296e-12          4.156218e-12          6.193690e-12
## Full Remote      5.058153e-12          1.338059e-12          5.077995e-12
##
## Residual Deviance: 712.461
## AIC: 772.461

## Likelihood ratio tests of Multinomial Models
##
## Response: remote_ratio
##
## 1                                                                 job_title + company_location
## 2 experience_level + employment_type + job_title + salary_in_usd + employee_residence + company_location
##   Resid. df Resid. Dev   Test      Df LR stat.    Pr(Chi)
## 1         864    693.6361
## 2         846    671.1218 1 vs 2     18  22.5143 0.2099499

## Call:
## multinom(formula = remote_ratio ~ job_title + company_location +
##   company_size, data = train, trace = FALSE)
##
## Coefficients:
##          (Intercept) job_titleData Scientist job_titleData Analyst
## Hybrid          0.06413182          0.5481136          0.7473942
## Full Remote      0.31661496          -0.1894828          0.4481260
##          job_titleML Engineer company_locationEU company_locationNA
## Hybrid          1.3390607          0.7554523          -0.7168916
## Full Remote      -0.1529826          0.2262414          0.9864969
##          company_sizeM company_sizeS
## Hybrid          -2.03168063          -0.4632989
## Full Remote      -0.09551458          0.5419590
##
## Std. Errors:
##          (Intercept) job_titleData Scientist job_titleData Analyst

```

```
## Hybrid          0.5980496          0.4399003          0.5636573
## Full Remote    0.4925914          0.2892217          0.3529384
##               job_titleML Engineer company_locationEU company_locationNA
## Hybrid          0.5747712          0.5510674          0.5419637
## Full Remote    0.4587180          0.4899853          0.4491222
##               company_sizeM company_sizeS
## Hybrid          0.4062267          0.5438868
## Full Remote    0.2818427          0.4664760
##
## Residual Deviance: 693.6361
## AIC: 725.6361
```

```
##               Reference
## Prediction      No Remote Hybrid Full Remote
## No Remote          0         0         0
## Hybrid             10        36        22
## Full Remote        84        35       253
```

```
## Accuracy
## 0.6568182
```

```
##               Sensitivity Specificity
## Class: No Remote    0.0000000    1.0000000
## Class: Hybrid       0.5070423    0.9132791
## Class: Full Remote  0.9200000    0.2787879
```

```
## Setting levels: control = FALSE, case = TRUE
```

```
## Setting direction: controls < cases
```

```
##
```

```
## Call:
```

```
## roc.default(response = (train$remote_ratio == "No Remote"), predictor = predprobs[,      1], percent = T
##
## Data: predprobs[, 1] in 346 controls ((train$remote_ratio == "No Remote") FALSE) < 94 cases ((train$rem
## Area under the curve: 62.97%
```

```
## Setting levels: control = FALSE, case = TRUE
```

```
## Setting direction: controls < cases
```

```
##
```

```
## Call:
```

```
## roc.default(response = (train$remote_ratio == "Hybrid"), predictor = predprobs[,      2], percent = T
##
## Data: predprobs[, 2] in 369 controls ((train$remote_ratio == "Hybrid") FALSE) < 71 cases ((train$rem
## Area under the curve: 87.24%
```

```
## Setting levels: control = FALSE, case = TRUE
```

```
## Setting direction: controls < cases
```

```
##
## Call:
## roc.default(response = (train$remote_ratio == "Full Remote"),      predictor = predprobs[, 3], percent
##
## Data: predprobs[, 3] in 165 controls ((train$remote_ratio == "Full Remote") FALSE) < 275 cases ((tra
## Area under the curve: 71.99%
```

