

# Freelancers

Iman

```
MEFreelancers <- read_csv("MiddleEastFreelancers.csv", show_col_types = FALSE)
```

```
str(MEFreelancers)
```

```
## spc_tbl_ [1,400 x 8] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ name      : chr [1:1400] "Ahmes.S" "Hossam A." "Mahmoud G." "Mustafa E." ...
## $ location   : chr [1:1400] "Egypt" "Egypt" "Egypt" "Egypt" ...
## $ price_per_hour: chr [1:1400] "$10" "$25" "$20" "$25" ...
## $ user_score  : num [1:1400] 24 24 24 24 24 24 24 24 24 24 ...
## $ verified_user : num [1:1400] 1 1 1 1 1 1 1 1 1 1 ...
## $ skills      : chr [1:1400] "Sustainable Energy" "2D Design,3D Design,3D Modeling,3D Rendering"
## $ description : chr [1:1400] NA "Interior, exterior, landscape and furniture designer" NA "Full s
## $ recommended : num [1:1400] 0 0 0 0 0 0 0 0 0 0 ...
## - attr(*, "spec")=
## .. cols(
## ..   name = col_character(),
## ..   location = col_character(),
## ..   price_per_hour = col_character(),
## ..   user_score = col_double(),
## ..   verified_user = col_double(),
## ..   skills = col_character(),
## ..   description = col_character(),
## ..   recommended = col_double()
## .. )
## - attr(*, "problems")=<externalptr>
```

```
head(MEFreelancers)
```

```
## # A tibble: 6 x 8
##   name      location price_per_hour user_score verified_user skills description
##   <chr>      <chr>      <chr>          <dbl>         <dbl> <chr>   <chr>
## 1 Ahmes.S    Egypt      $10              24             1 Susta~ <NA>
## 2 Hossam A.  Egypt      $25              24             1 2D De~ Interior, ~
## 3 Mahmoud G. Egypt      $20              24             1 Video~ <NA>
## 4 Mustafa E. Egypt      $25              24             1 .NET ~ Full stack~
## 5 Nader A.   Egypt      $8               24             1 Acade~ I`m Here t~
## 6 Ahmed T.   Egypt      $8               24             1 Medic~ general su~
## # i 1 more variable: recommended <dbl>
```

## Selecting Skills and Description

```
df_MEFreelancers <- MEFreelancers %>%
  select(skills, description)
head(df_MEFreelancers)
```

```
## # A tibble: 6 x 2
##   skills                                description
##   <chr>                                <chr>
## 1 Sustainable Energy                  <NA>
## 2 2D Design,3D Design,3D Modeling,3D Rendering Interior, ~
## 3 Video Production,English Translation,Information Management,IT Ma~ <NA>
## 4 .NET Compact Framework,.NET Development,Android App Development,A~ Full stack~
## 5 Academic Writing,Ad Design,Advertising,Art Curation             I`m Here t~
## 6 Medical Transcription,Medical Translation,Medical Writing        general su~
```

### Clean location Column

```
str(MEFreelancers$location)
```

```
## chr [1:1400] "Egypt" "Egypt" "Egypt" "Egypt" "Egypt" "Egypt" "Egypt" "Egypt" ...
```

```
# Filter out Israel
```

```
MEFreelancers <- MEFreelancers %>%
  filter(location != "Israel")
```

### Clean price\_per\_hour Column and Convert to OMR

```
MEFreelancers <- MEFreelancers %>%
  mutate(price_per_hour = as.numeric(gsub("\\$", "", price_per_hour)),
         price_per_hr_OR = price_per_hour * 0.38)
```

```
head(MEFreelancers)
```

```
## # A tibble: 6 x 9
##   name      location price_per_hour user_score verified_user skills description
##   <chr>      <chr>          <dbl>      <dbl>          <dbl> <chr> <chr>
## 1 Ahmes.S   Egypt              10          24              1 Susta~ <NA>
## 2 Hossam A. Egypt              25          24              1 2D De~ Interior, ~
## 3 Mahmoud G. Egypt              20          24              1 Video~ <NA>
## 4 Mustafa E. Egypt              25          24              1 .NET ~ Full stack~
## 5 Nader A.   Egypt              8           24              1 Acade~ I`m Here t~
## 6 Ahmed T.   Egypt              8           24              1 Medic~ general su~
## # i 2 more variables: recommended <dbl>, price_per_hr_OR <dbl>
```

### Count Freelancers by Location

```
MEFreelancers %>%
  count(location)
```

```
## # A tibble: 19 x 2
##   location      n
##   <chr>      <int>
## 1 Algeria        2
## 2 Bahrain        2
## 3 Egypt       185
## 4 India         88
## 5 Iran           4
## 6 Iraq           3
## 7 Jordan        31
## 8 Kuwait         5
```

```
## 9 Lebanon 72
## 10 Morocco 1
## 11 Oman 1
## 12 Pakistan 303
## 13 Palestine 34
## 14 Saudi Arabia 92
## 15 Sudan 1
## 16 Syria 1
## 17 Tunisia 27
## 18 Turkey 9
## 19 United Arab Emirates 537
```

### Count Recommended Freelancers by Location

```
MEFreelancers %>%
  filter(recommended == 1) %>%
  count(location)
```

```
## # A tibble: 5 x 2
##   location      n
##   <chr>      <int>
## 1 Egypt      3
## 2 Lebanon    4
## 3 Pakistan   2
## 4 Palestine   1
## 5 United Arab Emirates 8
```

### Price Summary (Min/Max)

```
MEFreelancers %>%
  summarize(min_price = min(price_per_hr_OR, na.rm = TRUE),
            max_price = max(price_per_hr_OR, na.rm = TRUE))
```

```
## # A tibble: 1 x 2
##   min_price max_price
##   <dbl>    <dbl>
## 1     0.38     380
```

### Price Summary by Location

```
MEFreelancers %>%
  group_by(location) %>%
  summarize(min_price = min(price_per_hr_OR, na.rm = TRUE),
            max_price = max(price_per_hr_OR, na.rm = TRUE))
```

```
## # A tibble: 19 x 3
##   location      min_price max_price
##   <chr>      <dbl>    <dbl>
## 1 Algeria      2.28      3.8
## 2 Bahrain      1.14      19
## 3 Egypt        0.38     380
## 4 India        0.38     22.8
## 5 Iran         1.52      95
## 6 Iraq         2.66      3.8
```

##	7	Jordan	1.14	57
##	8	Kuwait	5.7	19
##	9	Lebanon	1.14	38
##	10	Morocco	3.8	3.8
##	11	Oman	9.5	9.5
##	12	Pakistan	0.38	32.3
##	13	Palestine	1.9	13.3
##	14	Saudi Arabia	1.14	76
##	15	Sudan	3.8	3.8
##	16	Syria	1.9	1.9
##	17	Tunisia	0.38	19
##	18	Turkey	3.8	64.6
##	19	United Arab Emirates	0.76	380

Histogram of Recommended Status

```
ggplot(MEFreelancers, aes(x = factor(recommended))) +
  geom_bar(fill = "pink", color = "blue") +
  labs(title = "Histogram of Recommended Freelancers",
       x = "Recommended",
       y = "Count")
```

