

H2A Labor Insights

Scott Parisi

1/26/2022

The following Case Study uses fictitious names and businesses, but actual US Department of Labor Data.

The Business Challenge:

As an agricultural business in the United States, *East Granite Farms*, located in the Northeast USA, struggles with increasing challenges to stay profitable. They have determined their biggest challenge for the coming season is in finding enough seasonal labor to plant, harvest, and process their farm products. This challenge is shared with the entire agriculture industry as a whole and potentially affects the entire world labor force in the wake of the COVID-19 pandemic and the supply chain and labor issues it has brought forward. With the possibility of future crops not being planted or harvested which potentially could shudder a local farm, in operation for generations, *East Granite Farms* would like to apply for H2A seasonal migrant workers. They are unsure what wages to offer and how long of an engagement they should expect from the typical migrant worker, but desperately want to find an option to keep the farm operating.

George Houton, the Chief Operating Officer for *East Granite Farms*, recognizes the need for more information about the US H2A Visa program and how other business approach the application process. He has tasked the Data Analytics team to find relevant data and provide him some key insights into how to best assure that will be successful in their application for H2A Visa migrant workers, ensuring a successful season for the farm.

Step 1: Collecting Data

The Data Analytics team found a relevant data source through the US Department of Labor, accessible via <https://www.dol.gov/agencies/eta/foreign-labor/performance>.

This data source is from 2021 Q4, which represents H2A applications between *October 1st, 2020 & September 30th, 2021*, the most recent data available.

Step 2: Prepare the Data

The data was then prepared using Power BI, Power Query and DAX to create the following Measures:

Average Wage =

```
CALCULATE(  
    AVERAGE(  
        [WAGE_OFFER]  
    ),  
    'H-2A_Disclosure_FY2021Q4'[PER] = "Hour"  
)
```

Average Anticipated Hours =

```
AVERAGE(  
    'H-2A_Disclosure_FY2021Q4'[ANTICIPATED_NUMBER_OF_HOURS]  
)
```

```

Average Weeks of Employment =
DATEDIFF(
  AVERAGE(
    'H-2A_Disclosure_FY2021Q4' [EMPLOYMENT_BEGIN_DATE]
  ),
  AVERAGE(
    'H-2A_Disclosure_FY2021Q4' [EMPLOYMENT_END_DATE]
  ),
  WEEK
)

```

```

Average Total Compensation =
DATEDIFF(
  AVERAGE(
    'H-2A_Disclosure_FY2021Q4' [EMPLOYMENT_BEGIN_DATE]
  ),
  AVERAGE(
    'H-2A_Disclosure_FY2021Q4' [EMPLOYMENT_END_DATE]
  ),
  WEEK
) *
'H-2A_Disclosure_FY2021Q4' [Average Anticipated Hours]
*
'H-2A_Disclosure_FY2021Q4' [Average Wage]

```

These calculations create the necessary measurements and variables to summarize the compensation and length of visa at a state level.

Step 3: Process the data

The data was then grouped into a table by State using the following DAX:

```

Averageh2a =
SUMMARIZE(
  FILTER(
    'H-2A_Disclosure_FY2021Q4',
    [EMPLOYER_STATE] <> "PR" && [EMPLOYER_STATE] <> BLANK() &&
    [EMPLOYER_STATE] <> "MP" && [EMPLOYER_STATE] <> "VI"),
    [EMPLOYER_STATE],
    'Regions'[Region],
    "Average_Weeks_of_Employment", [Average Weeks of Employment],
    "Average_Anticipated_Hours", [Average Anticipated Hours],
    "Average_Total_Compensation", [Average Total Compensation],
    "Average_Wage", [Average Wage]
)

```

After exporting the table as a .CSV, the data is loaded into R Studio using the following code:

```

install.packages("tidyverse")

## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.1'
## (as 'lib' is unspecified)

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr  0.3.4

```

```
## v tibble 3.1.6      v dplyr 1.0.7
## v tidyr  1.1.4      v stringr 1.4.0
## v readr  2.1.1      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

h2a_full <- read_csv("H-2A_Disclosure_Data_FY2021_Q4_Summary.csv")

## Rows: 50 Columns: 6

## -- Column specification -----
## Delimiter: ","
## chr (2): EMPLOYER_STATE, Region
## dbl (4): Average_Weeks_of_Employment, Average_Total_Compensation, Average_Wa...

##
## 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.
```

Importing the data into R Studio provides an extensive platform for final analysis, visualization and reporting.

Step 4: Analyzing the Data

Considering the location of *East Granite Farms*, analysis was performed on the dataset to find averages and aggregations by region and state. Using the following formulas, information was gathered that was relevant to the business question at hand, **How can we use data-driven decision making to ensure enough labor is available for a prosperous harvest?**

Formula #1: Create a filtered data frame of just the Northeast States

```
h2a_northeast <- filter(h2a_full, Region == "Northeast")
```

Formula #2: Calculate Average Wage for the Northeast

```
avg_wage_northeast <- mean(h2a_northeast$Average_Wage)
```

Formula #3: Calculate Average Compensation for the Northeast

```
avg_compensation <- mean(h2a_full$Average_Total_Compensation)
```

Step 5: Visualize the Data

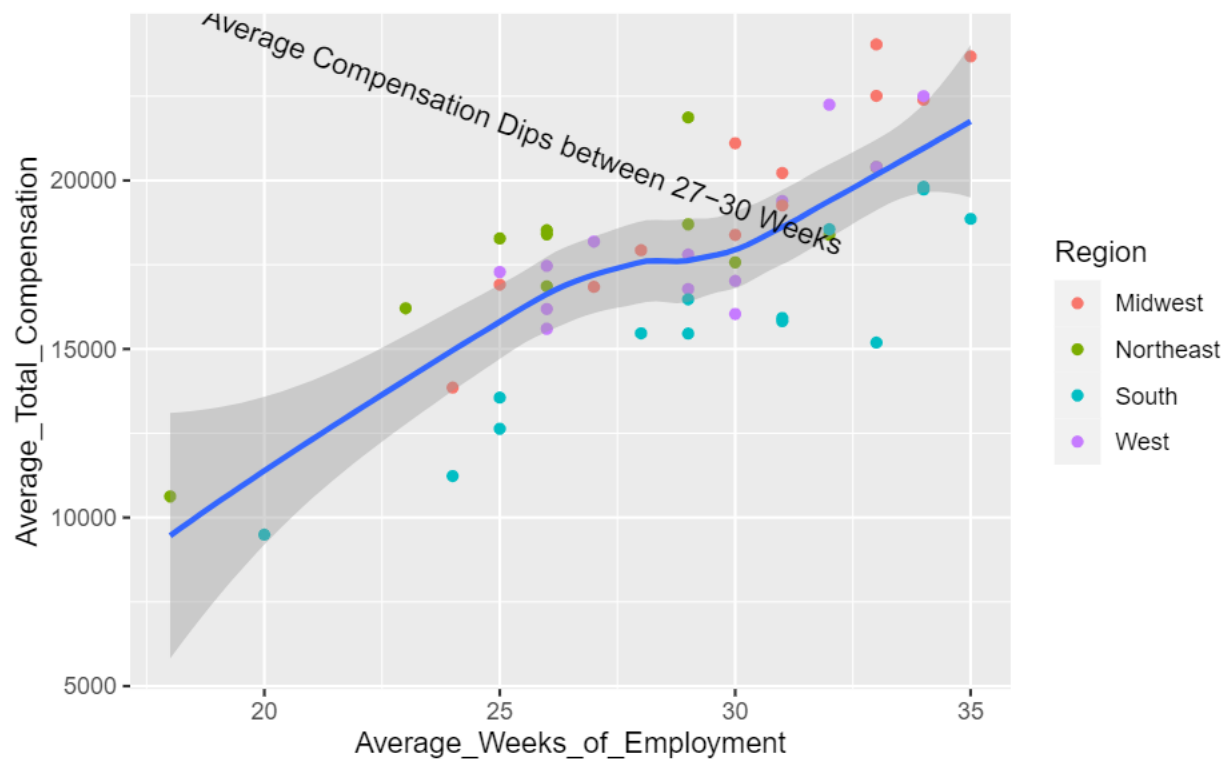
This visualization shows the intersection of the Average Total Compensation and Average Weeks of Employment per State.

This will help the COO and other stakeholders to see the national trends of total expense and time investment needed for the average H2A worker.

A trend in the visualization shows that it may be in the best interest of the company to not request an H2A worker longer than 30 weeks. Workers who are requested longer than 30 weeks need more investment in overall compensation, which is at a higher rate than in previous weeks.

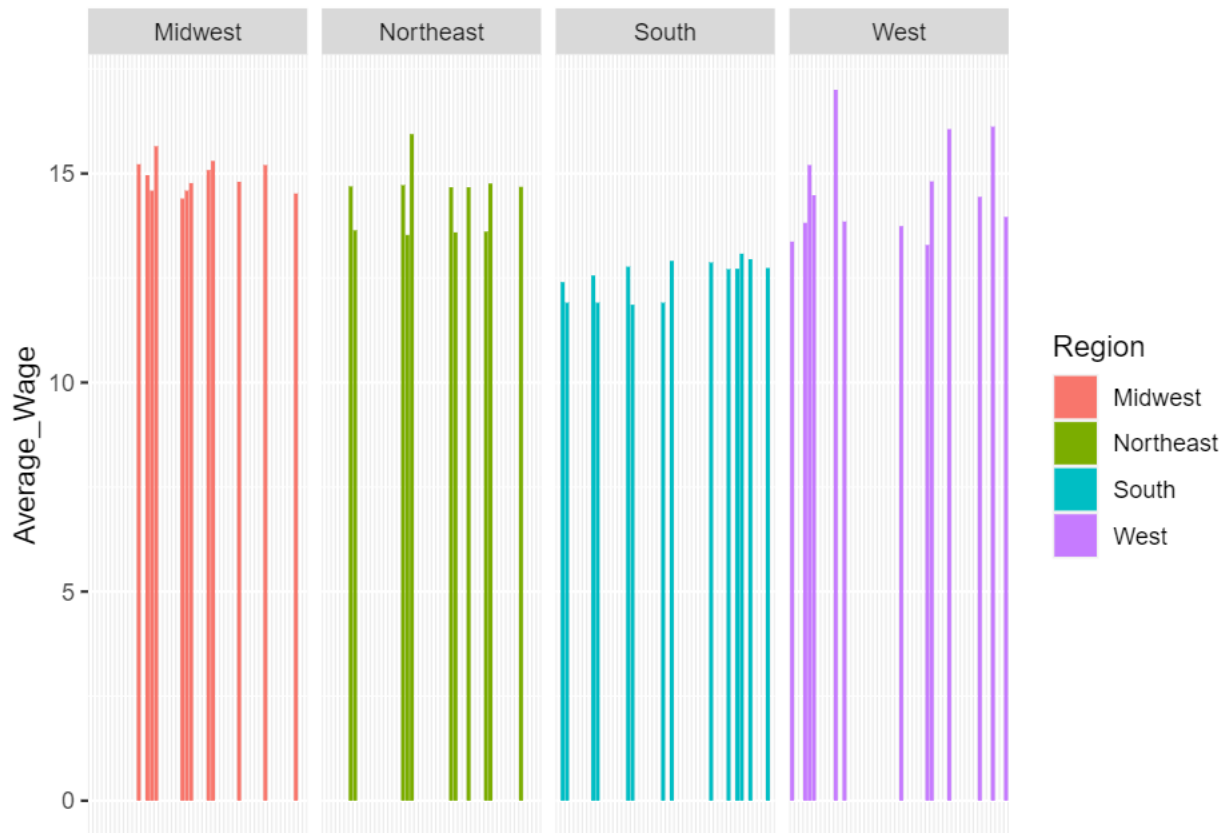
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

Average Total Compensation & Weeks of Employment Plotted by State of Employer



This visualization shows the average hourly wage paid to workers by region. Each bar represents a unique state. A trend shown here is that certain regions have higher or lower hourly wages, with the Northeast not the highest or the lowest, but in the middle.

This will help the COO and other stakeholders see the hourly wages they will need to pay for their region, and compare it to other regions in the country.



Step 6: Act on the Data

This analysis and visualization project has shown our stakeholders the average investment in time and dollars they will need to make in order to secure H2A migrant labor for the coming season. Beyond the obvious data, some high-level observations include a general lower investment in total compensation in southern states, a general higher investment in the west and northeast, and a need for higher pay for the location of *East Granite Farms*. This analysis will help the stakeholders of the farm make better decisions on labor in order to assure a bountiful harvest for the company and their community. Looking at it from the other perspective of the H2A laborer, they should enjoy higher compensation for the same amount of time as compared to working in the South region.