

Assignment 2

Deconstruct, Reconstruct Web Report

Michael Teixeira (S4133975)

Assessment declaration checklist

Please carefully read the statements below and check each box if you agree with the declaration. If you do not check all boxes, your assignment will not be marked. If you make a false declaration on any of these points, you may be investigated for academic misconduct. Students found to have breached academic integrity may receive official warnings and/or serious academic penalties. Please read more about academic integrity here (<https://www.rmit.edu.au/students/student-essentials/assessment-and-exams/academic-integrity>). If you are unsure about any of these points or feel your assessment might breach academic integrity, please contact your course coordinator for support. It is important that you DO NOT submit any assessment until you can complete the declaration truthfully.

By checking the boxes below, I declare the following:

- ☒ I have not impersonated, or allowed myself to be impersonated by, any person for the purposes of this assessment
- ☒ This assessment is my original work and no part of it has been copied from any other source except where due acknowledgement is made.
- ☒ No part of this assessment has been written for me by any other person except where such collaboration has been authorised by the lecturer/teacher concerned.
- ☒ Where this work is being submitted for individual assessment, I declare that it is my original work and that no part has been contributed by, produced by or in conjunction with another student.
- ☒ I give permission for my assessment response to be reproduced, communicated compared and archived for the purposes of detecting plagiarism.
- ☒ I give permission for a copy of my assessment to be retained by the university for review and comparison, including review by external examiners.

I understand that:

- ☒ Plagiarism is the presentation of the work, idea or creation of another person as though it is your own. It is a form of cheating and is a very serious academic offence that may lead to exclusion from the University. Plagiarised material can be drawn from, and presented in, written, graphic and visual form, including electronic data and oral presentations. Plagiarism occurs when the origin of the material used is not appropriately cited.
- ☒ Plagiarism includes the act of assisting or allowing another person to plagiarise or to copy my work.

I agree and acknowledge that:

- ☒ I have read and understood the Declaration and Statement of Authorship above.
- ☒ If I do not agree to the Declaration and Statement of Authorship in this context and all boxes are not checked, the assessment outcome is not valid for assessment purposes and will not be included in my final result for this course.

The visualisation chosen had the following three main issues:

- **Visual Complexity:** The circular/radial layout, while visually interesting, makes it difficult to compare values across occupations. The curved text and varying angles make it challenging to read occupation titles and compare growth percentages accurately.
- **Data Density Issues:** The visualization attempts to show too much information in a confined space, leading to overcrowded labels and potentially confusing overlaps between the salary figures and growth rates. The small icons, while adding visual interest, don't contribute meaningful information and add to the visual clutter.
- **Poor Use of Length Encoding:** The radial layout distorts the perception of the differences between values because the human eye is better at comparing lengths when they're aligned in parallel rather than at different angles. This makes it particularly difficult to compare salaries and growth rates across different occupations.

Reconstruct

Code

The following code was used to fix the issues identified in the original.

```

library(ggplot2)
library(dplyr)
library(tidyr)
library(scales)

# Create dataframe with the occupation data
occupations_data <- data.frame(
  occupation = c(
    "Wind turbine service technicians" = "Wind Turbine Tech.",
    "Solar photovoltaic installers" = "Solar PV Installers",
    "Nurse practitioners" = "Nurse Practitioners",
    "Data scientists" = "Data Scientists",
    "Information security analysts" = "Info. Security Analysts",
    "Medical and health services managers" = "Medical Services Mgrs.",
    "Physician assistants" = "Physician Assistants",
    "Computer and information research scientists" = "Computer Research Sci.",
    "Physical therapist assistants" = "Physical Therapy Asst.",
    "Operations research analysts" = "Operations Analysts",
    "Occupational therapy assistants" = "Occupational Therapy Asst.",
    "Actuaries" = "Actuaries",
    "Financial examiners" = "Financial Examiners",
    "Home health and personal care aides" = "Home Health Aides",
    "Veterinary assistants and laboratory animal caretakers" = "Vet. Lab Assistants",
    "Veterinary technologists and technicians" = "Vet. Technologists",
    "Logisticians" = "Logisticians",
    "Veterinarians" = "Veterinarians",
    "Substance abuse, behavioral disorder, and mental health counselors" = "Mental Health Cou
nseors",
    "Epidemiologists" = "Epidemiologists"
  ),
  growth_rate = c(60, 48, 46, 36, 33, 29, 28, 26, 25, 23, 22, 22, 21, 21, 19, 19, 19, 19, 19,
19),
  current_pay = c(61770, 48800, 126260, 108020, 120360, 110680, 130020, 145080, 64080, 83640,
67010, 120000, 84200, 33530, 36440, 43740, 79400, 119100, 53710, 81390)
)

# Calculate projected pay and format currency values
occupations_data <- occupations_data %>%
  mutate(
    projected_pay = current_pay * (1 + growth_rate / 100),
    current_pay_label = paste0("$", comma(round(current_pay/1000, 1)), "K"),
    projected_pay_label = paste0("$", comma(round(projected_pay/1000, 1)), "K")
  )

# Create the visualization
p1 <- ggplot(occupations_data, aes(x = reorder(occupation, -growth_rate))) +
  geom_segment(aes(y = current_pay, yend = projected_pay, xend = occupation, color = "Growth
Projection"),
    size = 0.8) +

  geom_point(aes(y = current_pay, color = "Current Pay"), size = 4) +
  geom_point(aes(y = projected_pay, color = "Projected Pay"), size = 4) +

  geom_text(aes(y = current_pay,
    label = sprintf("$.1fK", current_pay / 1000)),

```

```

    hjust = 1.5, size = 3, color = "#0072B2") +

geom_text(aes(y = projected_pay,
              label = sprintf("%.1fK", projected_pay / 1000)),
          hjust = -0.5, size = 3, color = "#009E73") +

geom_text(aes(y = max(current_pay, projected_pay) * 1.25,
              label = paste0(growth_rate, "%")),
          hjust = -0.2, size = 3.5, color = "gray40") +

# scale_y_continuous to expand the axis
scale_y_continuous(
  limits = c(0, max(occupations_data$projected_pay) * 1.3),
  breaks = seq(0, 200000, 50000),
  labels = function(x) paste0("$", x/1000, "K"),
  expand = expansion(mult = c(0.04, 0.1))
) +

coord_flip() +

labs(title = "Top U.S. Jobs by Projected Salary and Growth Rate (2023-2033)",
     subtitle = "Median Salary, Projected Salary, and Growth Rate for High-Demand Roles",
     x = "",
     y = "Annual Pay (Thousands of USD)",
     caption = "Data source: U.S. Bureau of Labor Statistics") +

theme_minimal() +
theme(
  plot.title = element_text(size = 14, face = "bold", margin = margin(t = 20, b = 10), hjust = 0.5),
  plot.subtitle = element_text(size = 12, margin = margin(b = 10), hjust = 0.5),
  axis.text.y = element_text(size = 10, color = "gray20", margin = margin(r = 5)),
  axis.text.x = element_text(size = 9, margin = margin(t = 0, b = 0)),
  panel.grid.major.y = element_blank(),
  panel.grid.minor = element_blank(),
  plot.margin = margin(t = 0, r = 70, b = 0, l = 0),
  plot.caption = element_text(color = "gray50", size = 9, hjust = 0.5),
  plot.background = element_rect(fill = "white", color = NA),
  panel.spacing.y = unit(1.5, "cm"),
  legend.position = "bottom"
) +

scale_x_discrete(expand = expansion(add = c(0.5, 0.5))) +

scale_color_manual(values = c("Current Pay" = "#0072B2", "Projected Pay" = "#009E73", "Growth Projection" = "gray70"),
                  name = "")

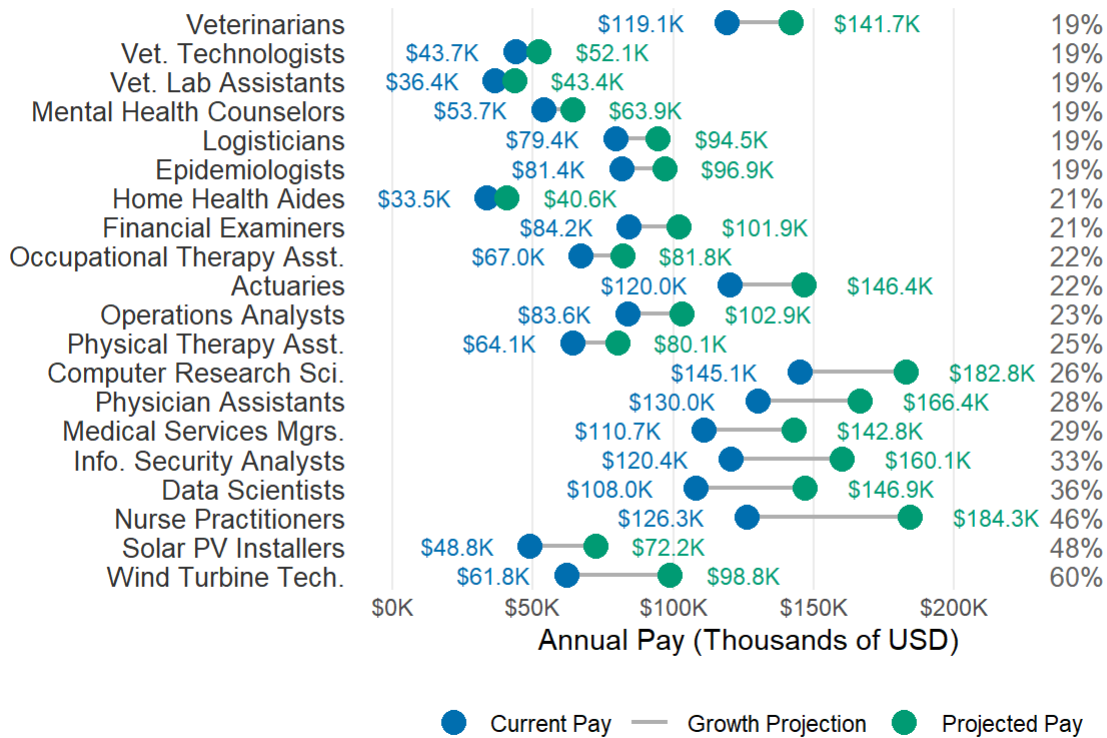
```

Reconstruction

The following plot fixes the main issues in the original.

Top U.S. Jobs by Projected Salary and Growth Rate (2023-2033)

Median Salary, Projected Salary, and Growth Rate for High-Demand Roles



Data source: U.S. Bureau of Labor Statistics

References

The reference to the original data visualisation choose, the data source(s) used for the reconstruction and any other sources used for this assignment are as follows:

- HowMuch.net. (2019). *Fastest Growing Occupations in the U.S.* Retrieved November 8, 2024, from HowMuch website: <https://howmuch.net/articles/fastest-growing-occupations-in-the-US> (<https://howmuch.net/articles/fastest-growing-occupations-in-the-US>)
- U.S. Bureau of Labor Statistics. (2023). *Fastest Growing Occupations.* Retrieved November 8, 2024, from U.S. Bureau of Labor Statistics website: <https://www.bls.gov/ooh/fastest-growing.htm> (<https://www.bls.gov/ooh/fastest-growing.htm>)