Highchart_Data_Viz

Pankaj Shah

6/30/2019

Contents

Libraries

```
library(tidyverse) #data Manipulation
library(highcharter) #interactive visualization
library(dplyr)
```

Reading Input dataset and Summary of it

```
survey18 <- read_csv("~/Downloads/stack-overflow-2018-developer-survey/survey_results_public.csv")
glimpse(survey18)
```

```
## Observations: 98,855
## Variables: 129
## $ Respondent                 <int> 1, 3, 4, 5, 7, 8, 9, 10, 11, 16, 1...
## $ Hobby                      <chr> "Yes", "Yes", "Yes", "No", "Yes", ...
## $ OpenSource                 <chr> "No", "Yes", "Yes", "No", "No", "N...
## $ Country                    <chr> "Kenya", "United Kingdom", "United...
## $ Student                    <chr> "No", "No", "No", "No", "Yes, part...
## $ Employment                 <chr> "Employed part-time", "Employed fu...
## $ FormalEducation            <chr> "Bachelor's degree (BA, BS, B.Eng....
## $ UndergradMajor             <chr> "Mathematics or statistics", "A na...
## $ CompanySize                <chr> "20 to 99 employees", "10,000 or m...
## $ DevType                    <chr> "Full-stack developer", "Database ...
## $ YearsCoding                <chr> "3-5 years", "30 or more years", "...
## $ YearsCodingProf            <chr> "3-5 years", "18-20 years", "6-8 y...
## $ JobSatisfaction            <chr> "Extremely satisfied", "Moderately...
## $ CareerSatisfaction         <chr> "Extremely satisfied", "Neither sa...
## $ HopeFiveYears              <chr> "Working as a founder or co-founde...
## $ JobSearchStatus            <chr> "I'm not actively looking, but I a...
## $ LastNewJob                 <chr> "Less than a year ago", "More than...
## $ AssessJob1                 <int> 10, 1, NA, NA, 8, 8, 5, 6, 6, NA, ...
## $ AssessJob2                 <int> 7, 7, NA, NA, 5, 5, 3, 5, 3, NA, N...
## $ AssessJob3                 <int> 8, 10, NA, NA, 7, 4, 9, 4, 7, NA, ...
## $ AssessJob4                 <int> 1, 8, NA, NA, 1, 9, 4, 2, 4, NA, N...
## $ AssessJob5                 <int> 2, 2, NA, NA, 2, 1, 1, 7, 1, NA, N...
## $ AssessJob6                 <int> 5, 5, NA, NA, 6, 3, 8, 8, 5, NA, N...
## $ AssessJob7                 <int> 3, 4, NA, NA, 4, 6, 2, 10, 10, NA,...
## $ AssessJob8                 <int> 4, 3, NA, NA, 3, 2, 7, 1, 8, NA, N...
## $ AssessJob9                 <int> 9, 6, NA, NA, 10, 10, 10, 9, 9, NA...
## $ AssessJob10                <int> 6, 9, NA, NA, 9, 7, 6, 3, 2, NA, N...
## $ AssessBenefits1            <int> NA, 1, NA, NA, 1, 1, 1, 1, 1, NA, ...
## $ AssessBenefits2            <int> NA, 5, NA, NA, 10, 3, 3, 3, 3, NA,...
## $ AssessBenefits3            <int> NA, 3, NA, NA, 2, 4, 2, 5, 2, NA, ...
## $ AssessBenefits4            <int> NA, 7, NA, NA, 4, 10, 9, 7, 9, NA,...
## $ AssessBenefits5            <int> NA, 10, NA, NA, 8, 9, 11, 6, 11, N...
## $ AssessBenefits6            <int> NA, 4, NA, NA, 3, 2, 4, 2, 5, NA, ...
## $ AssessBenefits7            <int> NA, 11, NA, NA, 11, 6, 8, 11, 8, N...
## $ AssessBenefits8            <int> NA, 9, NA, NA, 7, 5, 6, 9, 4, NA, ...
## $ AssessBenefits9            <int> NA, 6, NA, NA, 5, 11, 7, 4, 10, NA...
## $ AssessBenefits10           <int> NA, 2, NA, NA, 9, 7, 10, 10, 7, NA...
## $ AssessBenefits11           <int> NA, 8, NA, NA, 6, 8, 5, 8, 6, NA, ...
```

```
## $ JobContactPriorities1       <int> 3, 3, NA, NA, 2, 4, 3, 1, 5, NA, N...
## $ JobContactPriorities2       <int> 1, 1, NA, NA, 1, 2, 1, 3, 1, NA, N...
## $ JobContactPriorities3       <int> 4, 5, NA, NA, 4, 5, 5, 2, 2, NA, N...
## $ JobContactPriorities4       <int> 2, 2, NA, NA, 5, 1, 4, 4, 3, NA, N...
## $ JobContactPriorities5       <int> 5, 4, NA, NA, 3, 3, 2, 5, 4, NA, N...
## $ JobEmailPriorities1         <int> 5, 1, NA, NA, 7, 2, 1, 2, 3, NA, N...
## $ JobEmailPriorities2         <int> 6, 3, NA, NA, 3, 6, 5, 6, 7, NA, N...
## $ JobEmailPriorities3         <int> 7, 4, NA, NA, 6, 7, 3, 1, 2, NA, N...
## $ JobEmailPriorities4         <int> 2, 5, NA, NA, 2, 3, 4, 3, 4, NA, N...
## $ JobEmailPriorities5         <int> 1, 2, NA, NA, 1, 1, 2, 7, 1, NA, N...
## $ JobEmailPriorities6         <int> 4, 6, NA, NA, 4, 5, 6, 5, 6, NA, N...
## $ JobEmailPriorities7         <int> 3, 7, NA, NA, 5, 4, 7, 4, 5, NA, N...
## $ UpdateCV                    <chr> "My job status or other personal s...
## $ Currency                    <chr> NA, "British pounds sterling (£)",...
## $ Salary                      <dbl> NA, 51000, NA, NA, 260000, 30000, ...
## $ SalaryType                  <chr> "Monthly", "Yearly", NA, NA, "Year...
## $ ConvertedSalary             <dbl> NA, 70841, NA, NA, 21426, 41671, 1...
## $ CurrencySymbol              <chr> "KES", "GBP", NA, NA, "ZAR", "GBP"...
## $ CommunicationTools          <chr> "Slack", "Confluence;Office / prod...
## $ TimeFullyProductive         <chr> "One to three months", "One to thr...
## $ EducationTypes              <chr> "Taught yourself a new language, f...
## $ SelfTaughtTypes             <chr> "The official documentation and/or...
## $ TimeAfterBootcamp           <chr> NA, NA, NA, NA, NA, NA, NA, "Immed...
## $ HackathonReasons            <chr> "To build my professional network"...
## $ AgreeDisagree1              <chr> "Strongly agree", "Agree", NA, "Di...
## $ AgreeDisagree2              <chr> "Strongly agree", "Agree", NA, "Di...
## $ AgreeDisagree3              <chr> "Neither Agree nor Disagree", "Nei...
## $ LanguageWorkedWith          <chr> "JavaScript;Python;HTML;CSS", "Jav...
## $ LanguageDesireNextYear      <chr> "JavaScript;Python;HTML;CSS", "Go;...
## $ DatabaseWorkedWith          <chr> "Redis;SQL Server;MySQL;PostgreSQL...
## $ DatabaseDesireNextYear      <chr> "Redis;SQL Server;MySQL;PostgreSQL...
## $ PlatformWorkedWith          <chr> "AWS;Azure;Linux;Firebase", "Linux...
## $ PlatformDesireNextYear      <chr> "AWS;Azure;Linux;Firebase", "Linux...
## $ FrameworkWorkedWith         <chr> "Django;React", "Django", NA, NA, ...
## $ FrameworkDesireNextYear     <chr> "Django;React", "React", NA, "Angu...
## $ IDE                         <chr> "Komodo;Vim;Visual Studio Code", "...
## $ OperatingSystem             <chr> "Linux-based", "Linux-based", NA, ...
## $ NumberMonitors              <chr> "1", "2", NA, "2", "2", "2", "2", ...
## $ Methodology                 <chr> "Agile;Scrum", NA, NA, "Agile;Kanb...
## $ VersionControl              <chr> "Git", "Git;Subversion", NA, "Git"...
## $ CheckInCode                 <chr> "Multiple times per day", "A few t...
## $ AdBlocker                   <chr> "Yes", "Yes", NA, "Yes", "No", "Ye...
## $ AdBlockerDisable            <chr> "No", "Yes", NA, "Yes", NA, "Yes",...
## $ AdBlockerReasons            <chr> NA, "The website I was visiting as...
## $ AdsAgreeDisagree1           <chr> "Strongly agree", "Somewhat agree"...
## $ AdsAgreeDisagree2           <chr> "Strongly agree", "Neither agree n...
## $ AdsAgreeDisagree3           <chr> "Strongly agree", "Neither agree n...
## $ AdsActions                  <chr> "Saw an online advertisement and t...
## $ AdsPriorities1              <int> 1, 3, NA, NA, 2, 1, 1, NA, 1, NA, ...
## $ AdsPriorities2              <int> 5, 5, NA, NA, 3, 3, 4, NA, 3, NA, ...
## $ AdsPriorities3              <int> 4, 1, NA, NA, 4, 4, 2, NA, 5, NA, ...
## $ AdsPriorities4              <int> 7, 4, NA, NA, 6, 2, 5, NA, 4, NA, ...
## $ AdsPriorities5              <int> 2, 6, NA, NA, 1, 7, 3, NA, 2, NA, ...
## $ AdsPriorities6              <int> 6, 7, NA, NA, 7, 5, 7, NA, 7, NA, ...
```

```
## $ AdsPriorities7              <int> 3, 2, NA, NA, 5, 6, 6, NA, 6, NA, ...
## $ AIDangerous                 <chr> "Artificial intelligence surpassin...
## $ AIInteresting               <chr> "Algorithms making important decis...
## $ AIResponsible               <chr> "The developers or the people crea...
## $ AIFuture                    <chr> "I'm excited about the possibiliti...
## $ EthicsChoice                <chr> "No", "Depends on what it is", NA,...
## $ EthicsReport                <chr> "Yes, and publicly", "Depends on w...
## $ EthicsResponsible           <chr> "Upper management at the company/o...
## $ EthicalImplications         <chr> "Yes", "Yes", NA, "Yes", "Yes", "U...
## $ StackOverflowRecommend      <chr> "10 (Very Likely)", "10 (Very Like...
## $ StackOverflowVisit          <chr> "Multiple times per day", "A few t...
## $ StackOverflowHasAccount     <chr> "Yes", "Yes", NA, "Yes", "Yes", "Y...
## $ StackOverflowParticipate    <chr> "I have never participated in Q&A ...
## $ StackOverflowJobs           <chr> "No, I knew that Stack Overflow ha...
## $ StackOverflowDevStory       <chr> "Yes", "No, I have one but it's ou...
## $ StackOverflowJobsRecommend  <chr> NA, "7", NA, "8", NA, "8", "7", NA...
## $ StackOverflowConsiderMember <chr> "Yes", "Yes", NA, "Yes", "Yes", "N...
## $ HypotheticalTools1          <chr> "Extremely interested", "A little ...
## $ HypotheticalTools2          <chr> "Extremely interested", "A little ...
## $ HypotheticalTools3          <chr> "Extremely interested", "A little ...
## $ HypotheticalTools4          <chr> "Extremely interested", "A little ...
## $ HypotheticalTools5          <chr> "Extremely interested", "A little ...
## $ WakeTime                    <chr> "Between 5:00 - 6:00 AM", "Between...
## $ HoursComputer               <chr> "9 - 12 hours", "5 - 8 hours", NA,...
## $ HoursOutside                <chr> "1 - 2 hours", "30 - 59 minutes", ...
## $ SkipMeals                   <chr> "Never", "Never", NA, "3 - 4 times...
## $ ErgonomicDevices            <chr> "Standing desk", "Ergonomic keyboa...
## $ Exercise                    <chr> "3 - 4 times per week", "Daily or ...
## $ Gender                      <chr> "Male", "Male", NA, "Male", "Male"...
## $ SexualOrientation           <chr> "Straight or heterosexual", "Strai...
## $ EducationParents            <chr> "Bachelor's degree (BA, BS, B.Eng....
## $ RaceEthnicity               <chr> "Black or of African descent", "Wh...
## $ Age                         <chr> "25 - 34 years old", "35 - 44 year...
## $ Dependents                  <chr> "Yes", "Yes", NA, "No", "Yes", "No...
## $ MilitaryUS                  <chr> NA, NA, NA, "No", NA, NA, "No", NA...
## $ SurveyTooLong               <chr> "The survey was an appropriate len...
## $ SurveyEasy                  <chr> "Very easy", "Somewhat easy", NA, ...
```

`highcharter` lets you plot using two different functions:

- highchart()
- hchart()

highchart()

This function creates a Highchart chart using htmlwidgets. The widget can be rendered on HTML pages generated from R Markdown, Shiny, or other applications.Similarly, Once the `highchart()` function is defined further highchart elements can be added on top of it.

hchart()

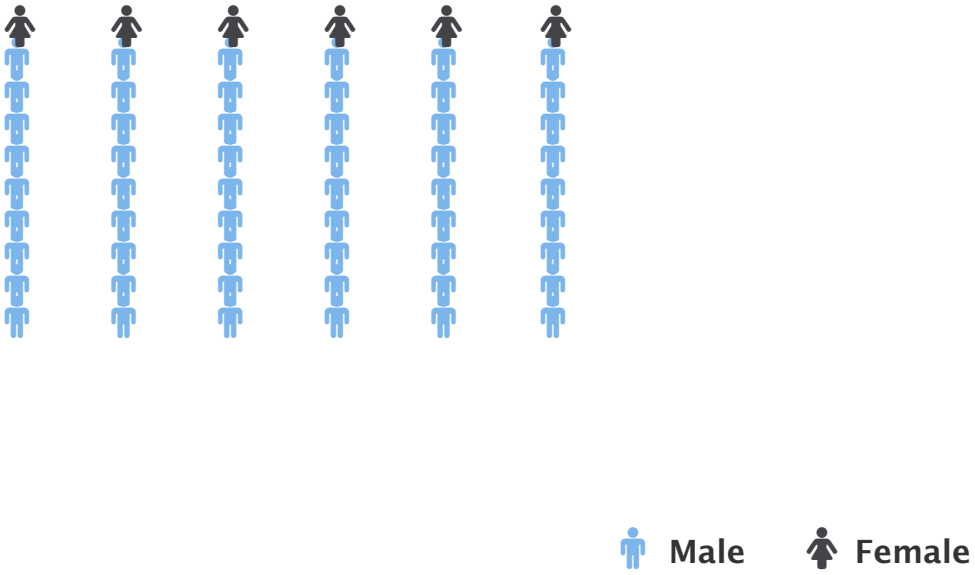
`hchart()` is a generic function to draw different charts on the fly. The resulting chart is a highchart object so you can keep modifying with the implmented API. If you are familiar with `ggplot2`, this function is similar

to `qplot()` of it.

Let us begin our Interactive Visualization journey with the easy plots.

Icons Plot

```
survey18 %>%
  dplyr::select(Gender) %>%
  filter(!is.na(Gender)) %>%
  filter(Gender %in% c('Male','Female')) %>%
  count(Gender) %>%
  mutate(perc = round((n /sum(n))*100)) -> gender_icons
  hciconarray(c('Female','Male'),gender_icons$perc,icons = c('male','female'))
```

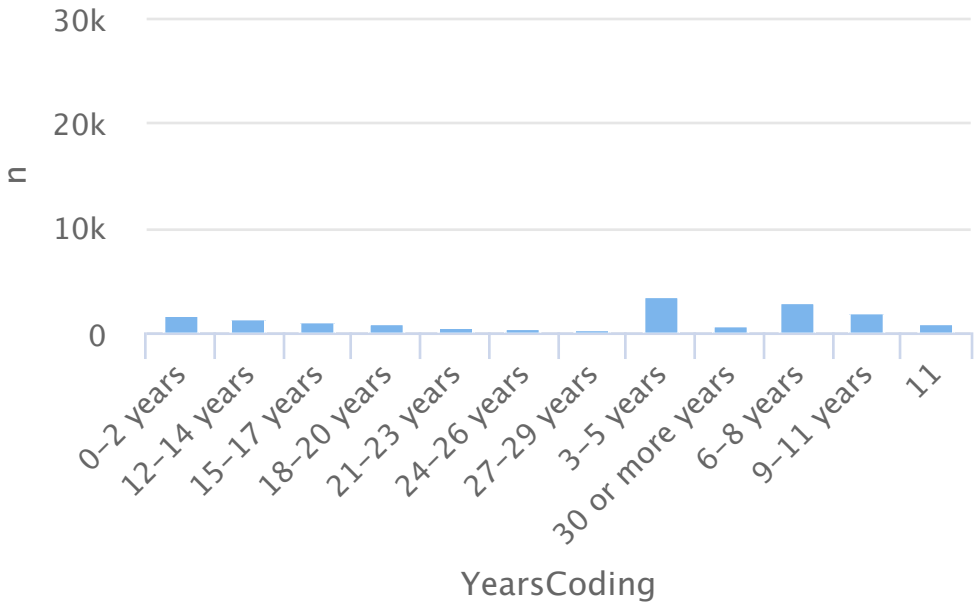


Bar

hchart - column

Barplot is useful when you have comparable Categorical variables (factors). Let us look at what Years of Coding experience the respondents have got.

```
survey18 %>%
  count(YearsCoding) %>%
  hchart('column', hcaes(x = 'YearsCoding', y = 'n'))
```

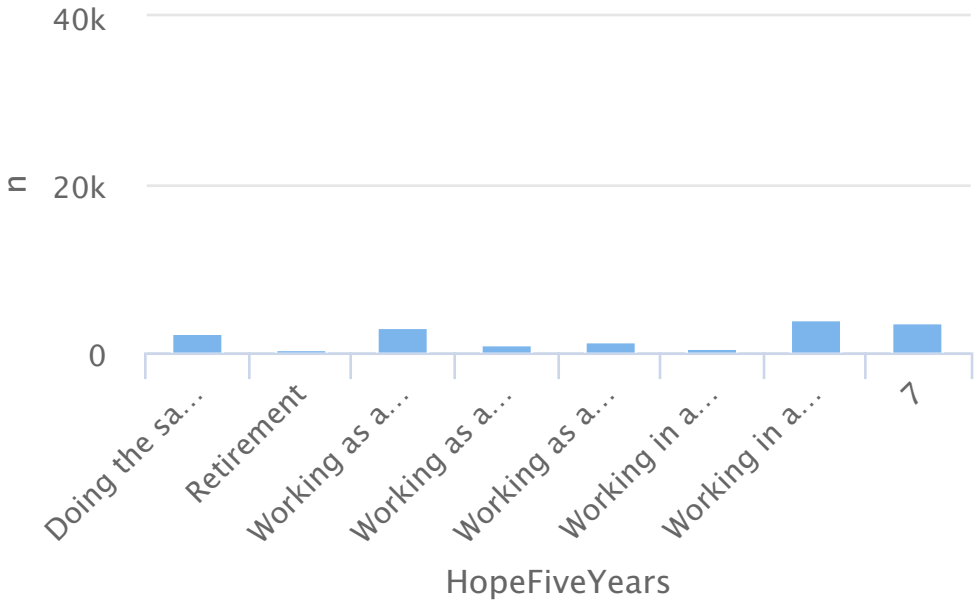


As you can see in the above code, the chart type here is `column` that makes a vertical bar plot. Aesthetics are given using `hcaes()` (similar to `aes()` of `ggplot2`).

hchart - bar

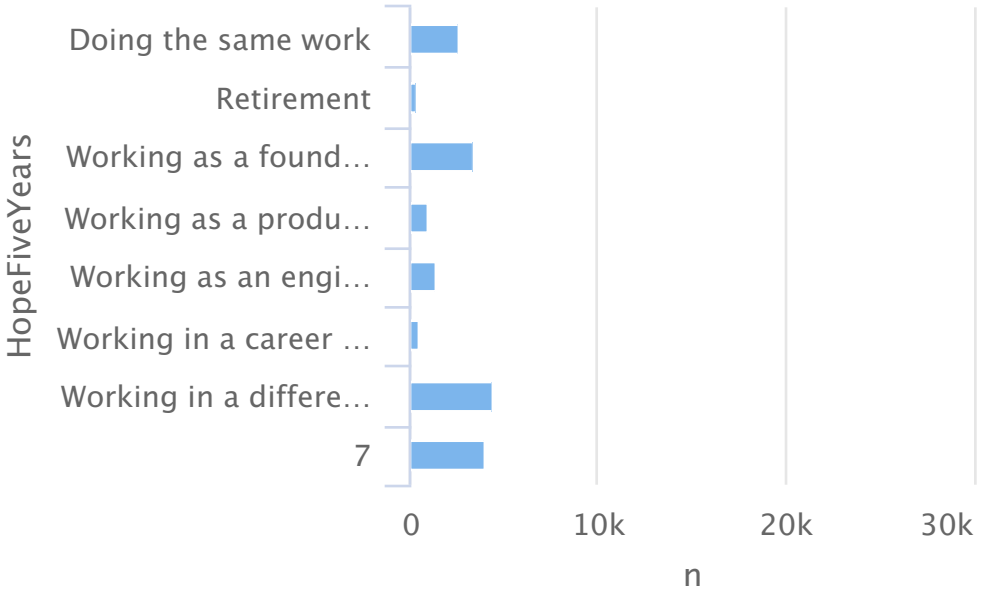
Let us look at what people are hoping for five years.

```
survey18 %>%
  count(HopeFiveYears) %>%
  hchart('column', hcaes(x = 'HopeFiveYears', y = 'n'))
```



while the same column plot as above does the job, it can seen the large axis label have to be rotated and also cut - which may not be something fine always. Hence we will rotate the plot to make it a horizontal bar plot so large axis label can be accomodated.

```
survey18 %>%
  count(HopeFiveYears) %>%
  hchart('bar', hcaes(x = 'HopeFiveYears', y = 'n'))
```

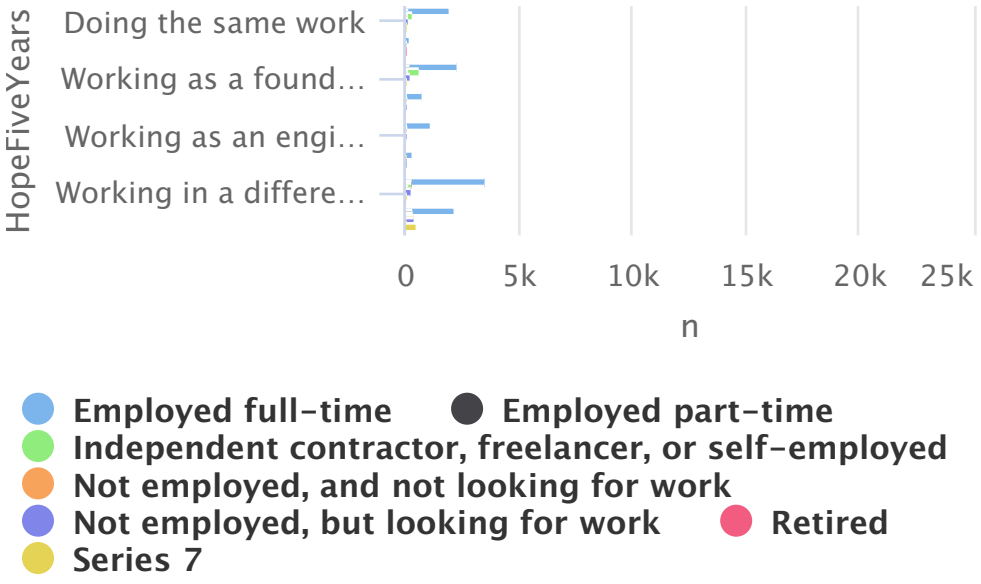


The chart type here is `bar`.

Grouped Bar

Let us try to add one more dimension to our existing bar, by seeing how this hope for next five years varies based on their employement type.

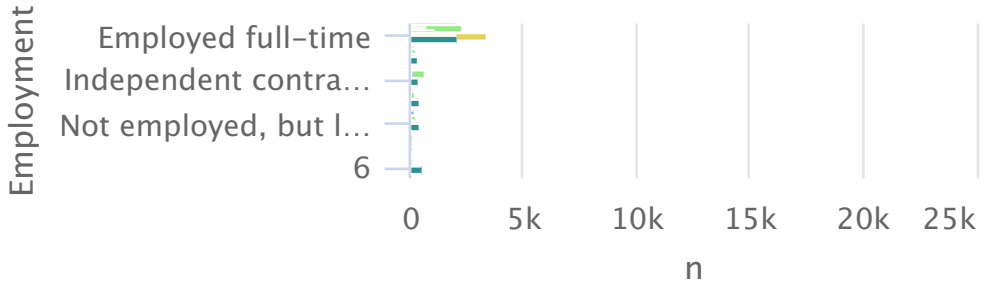
```
survey18 %>%
  count(Employment, HopeFiveYears) %>%
  hchart('bar', hcaes(x = 'HopeFiveYears', y = 'n', group = 'Employment'))
```



Now, this chart is as same as the above one except with the addition of grouped by Employement type.

But the grouping can be flipped to see the story from a different lens.

```
survey18 %>%
  count(Employment, HopeFiveYears) %>%
  hchart('bar', hcaes(x = 'Employment', y = 'n', group = 'HopeFiveYears'))
```



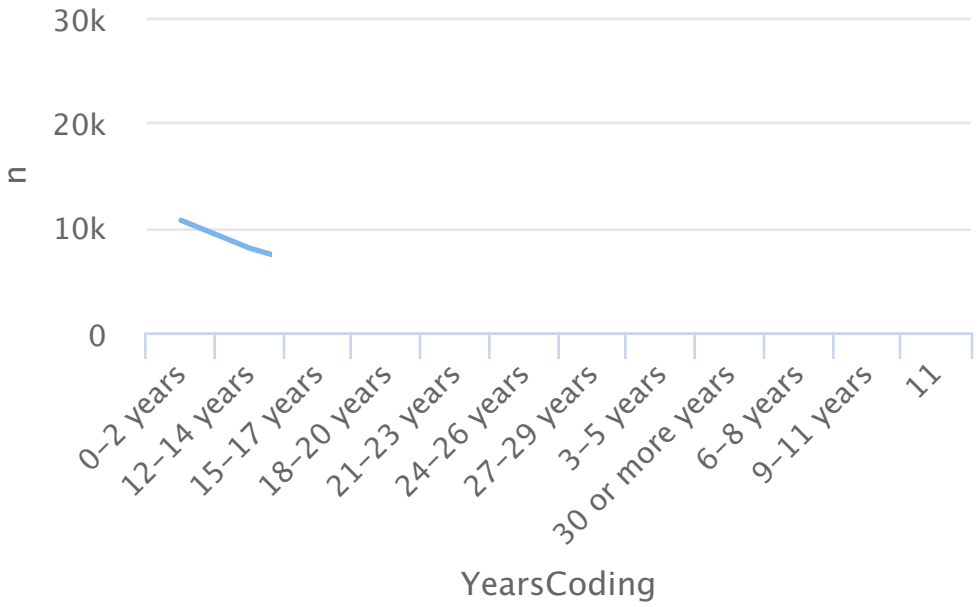
he same work ● Retirement
g as a founder or co-founder of my own company
g as a product manager or project manager
g as an engineering manager or other functional manager
g in a career completely unrelated to software development
g in a different or more specialized technical role than the one
3

Line & Area

Line

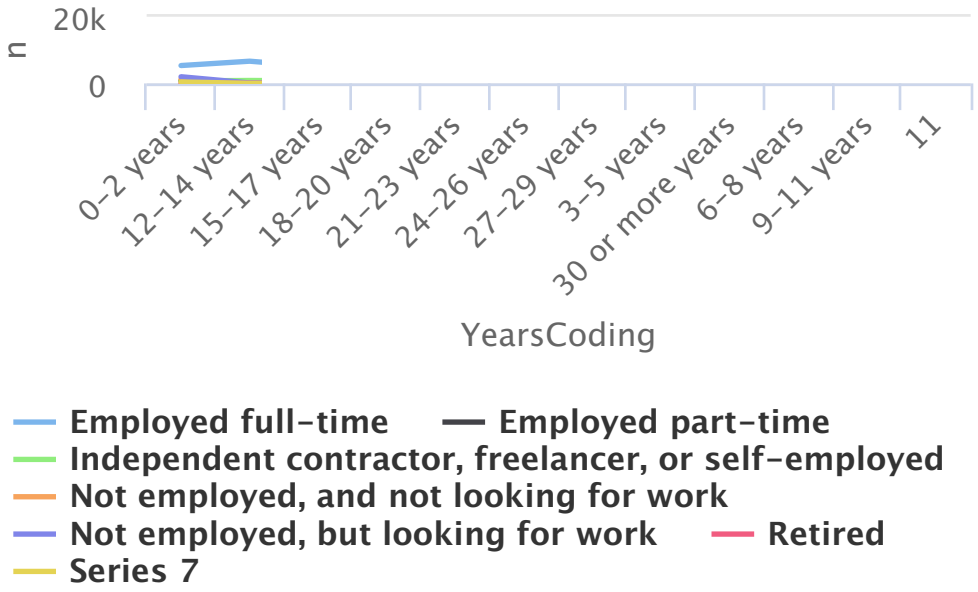
Line is particularly prefered when you have Time variable in x-axis but considering we don't have any Time variable in the given dataset, we can chart out categorical varaibles using Line graph.

```
survey18 %>%
  count(YearsCoding) %>%
  hchart('line', hcaes(x = 'YearsCoding', y = 'n'))
```



Grouped Line

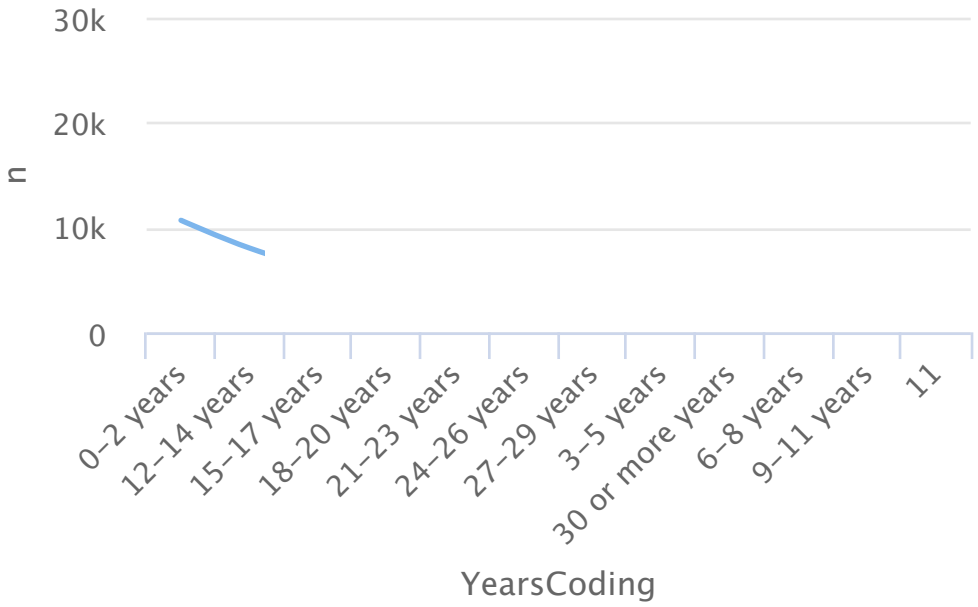
```
survey18 %>%
  count(YearsCoding, Employment) %>%
  hchart('line', hcaes(x = 'YearsCoding', y = 'n', group = "Employment"))
```



Spline

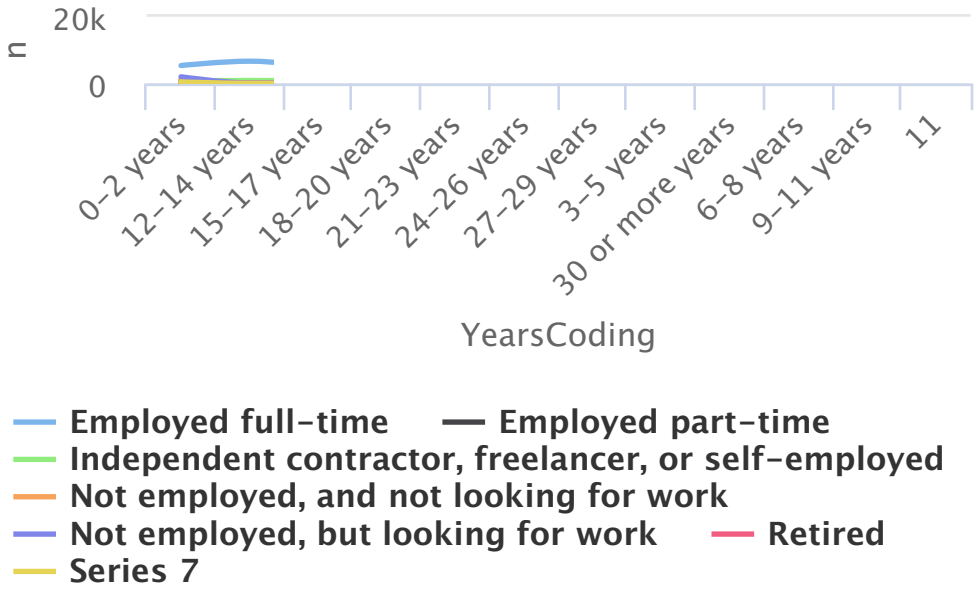
Line with Polynomial Interpolation.

```
survey18 %>%
  count(YearsCoding) %>%
  hchart('spline', hcaes(x = 'YearsCoding', y = 'n'))
```



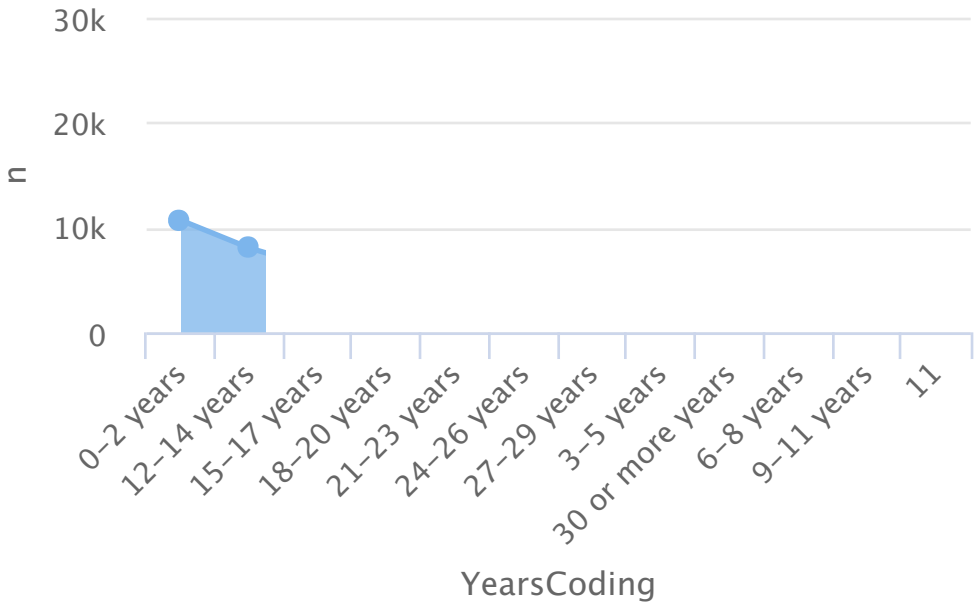
Grouped Spline

```
survey18 %>%
  count(YearsCoding, Employment) %>%
  hchart('spline', hcaes(x = 'YearsCoding', y = 'n', group = "Employment"))
```



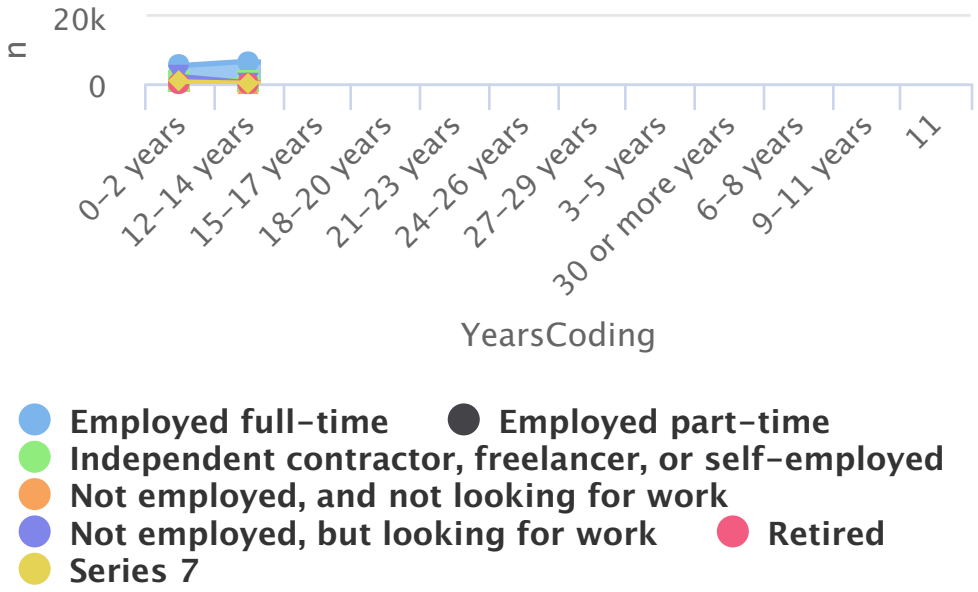
Area

```
survey18 %>%
  count(YearsCoding) %>%
  hchart('area', hcaes(x = 'YearsCoding', y = 'n'))
```



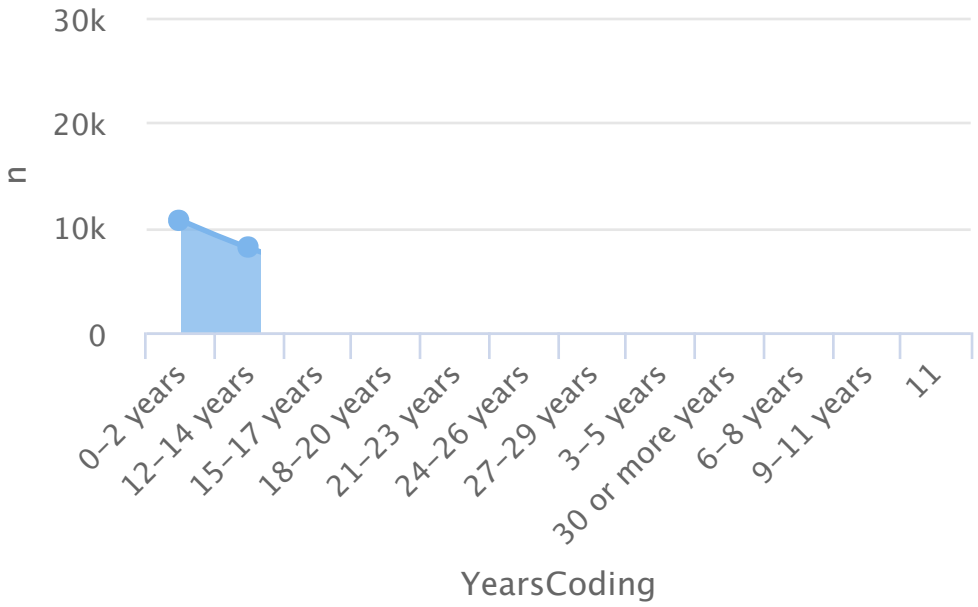
Grouped Area

```
survey18 %>%
  count(YearsCoding, Employment) %>%
  hchart('area', hcaes(x = 'YearsCoding', y = 'n', group = "Employment"))
```



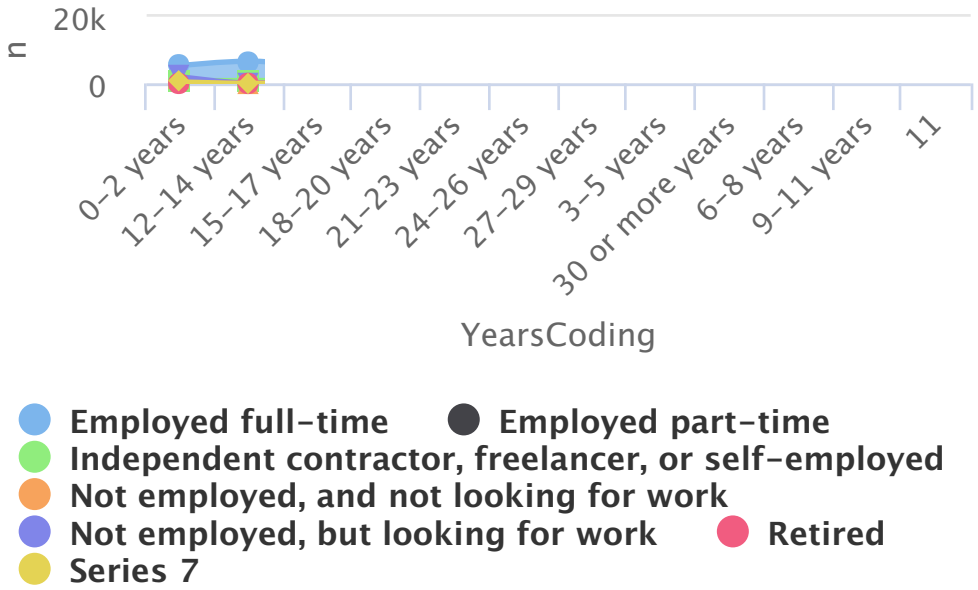
Area Spline

```
survey18 %>%
  count(YearsCoding) %>%
  hchart('areaspline', hcaes(x = 'YearsCoding', y = 'n'))
```



Grouped Area Spline

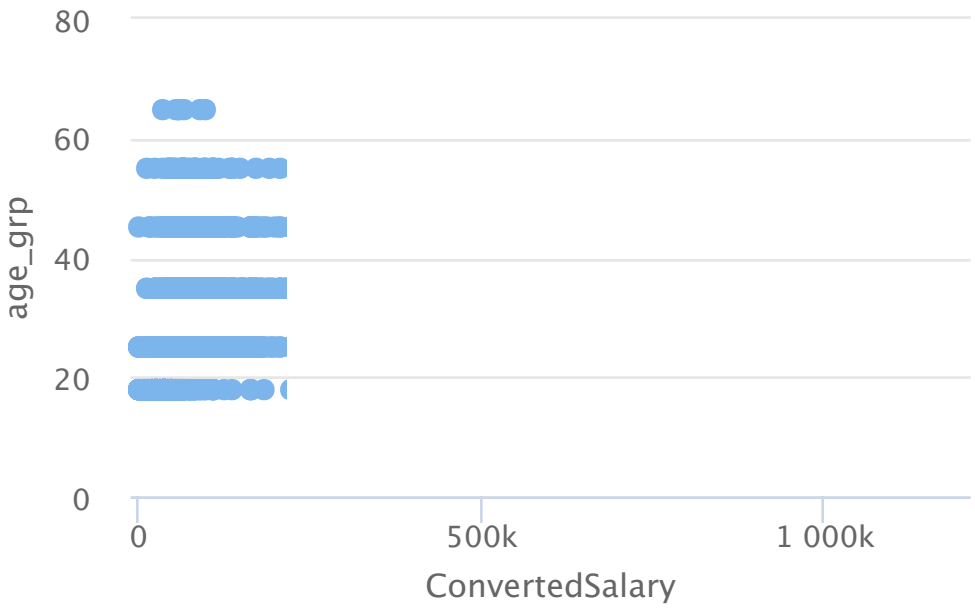
```
survey18 %>%
  count(YearsCoding, Employment) %>%
  hchart('areaspline', hcaes(x = 'YearsCoding', y = 'n', group = "Employment"))
```



As you can see in all the above plots, it's just the chart type changes and yet for the same data `hchart()` function is capable of managing to plot a different chart with the same data which makes this package really an easy way to make plots.

Scatter Plot

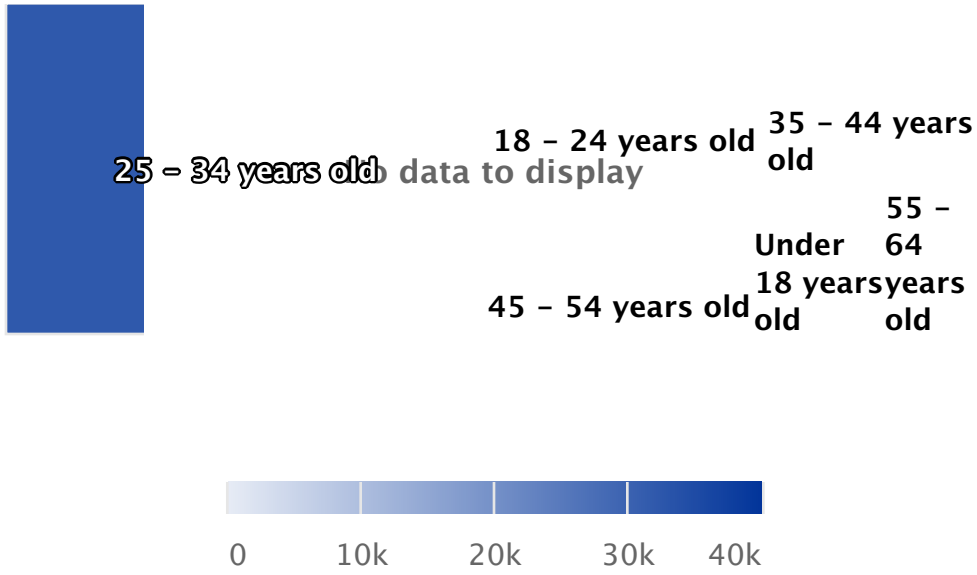
```
survey18 %>%
  filter(!is.na(Gender),
         Gender %in% c('Male','Female')) %>%
  filter(Country %in% 'United Kingdom') %>%
  filter(!is.na(Age),
         !is.na(ConvertedSalary),
         ConvertedSalary > 0) %>%
  dplyr::select(Gender, Country, Age, ConvertedSalary) %>%
  mutate(age_grp = parse_number(Age)) %>%
  hchart('scatter',hcaes('ConvertedSalary',"age_grp"))
```



Treemap

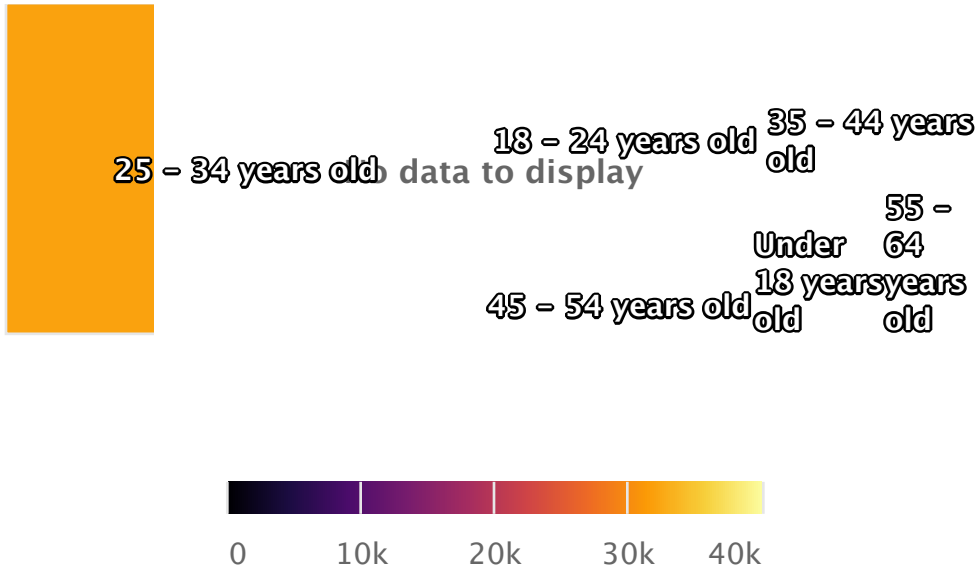
Boring

```
survey18 %>%
  filter(!is.na(Age)) %>%
  count(Age) %>%
  hchart('treemap', hcaes(x = 'Age', value = 'n', color = 'n'))
```



Less Boring

```
survey18 %>%
  filter(!is.na(Age)) %>%
  count(Age) %>%
  hchart('treemap', hcaes(x = 'Age', value = 'n', color = 'n')) %>%
    hc_colorAxis(stops = color_stops(colors = viridis::inferno(10)))
```



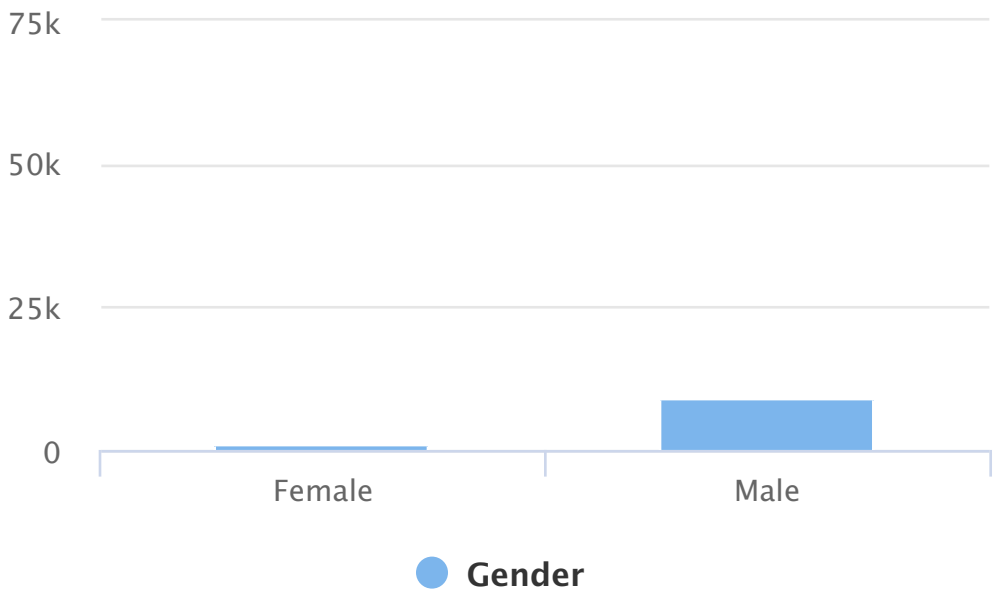
highchart() Type Charts

Bar

Most of the charts above can reproduced using the other method of `highchart()` and building layers on top of it. To start with, we will explain how we can build a bar chart like that.

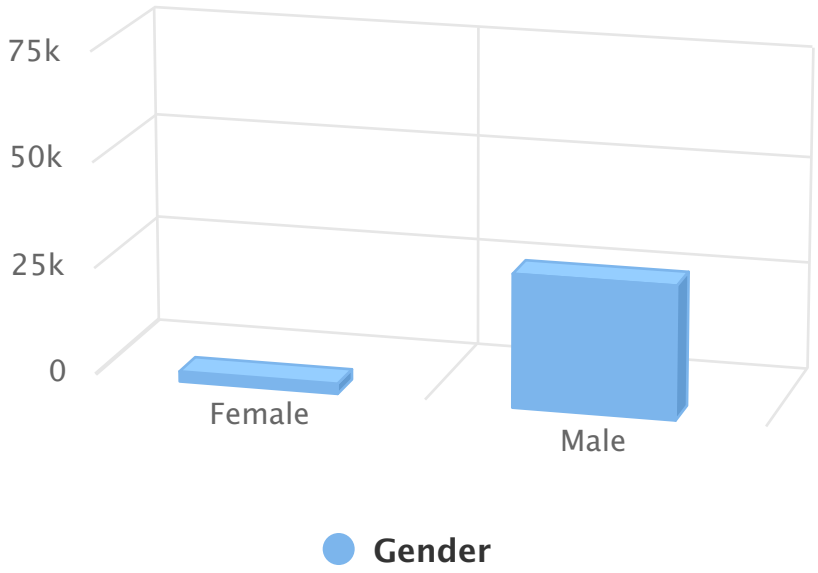
```
survey18 %>%
  filter(!is.na(Gender)) %>%
  filter(Gender %in% c('Male','Female')) %>%
  count(Gender) -> Gender

highchart() %>%
  hc_chart(type ="column") %>%
  hc_xAxis(categories = Gender$Gender) %>%
  hc_add_series(data = Gender$n, name = "Gender")
```



3D-Bar

```
highchart() %>%
  hc_chart(type ="column",
           options3d = list(enabled = TRUE, beta = 15, alpha = 15)) %>%
  hc_xAxis(categories = Gender$Gender) %>%
  hc_add_series(data = Gender$n, name = "Gender")
```



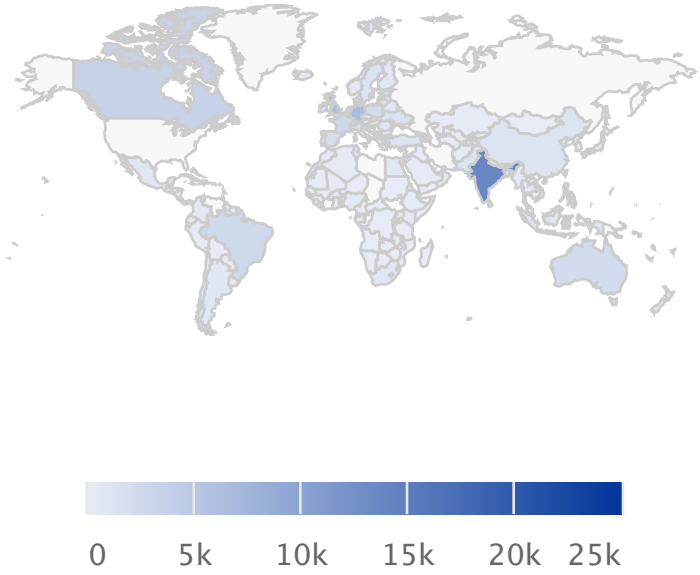
Maps

Nice but boring and incomplete

```
data(worldgeojson, package = "highcharter")

survey18 %>%
  filter(!is.na(Country)) %>%
  count(Country) -> countries

highchart() %>%
hc_add_series_map(worldgeojson, countries, value = "n", joinBy = c('name','Country'))
```

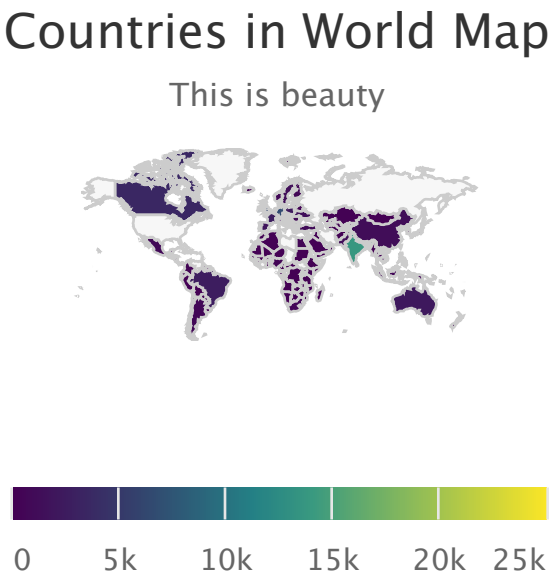


Colorful Map with Title and Subtitle - Complete

```
data(worldgeojson, package = "highcharter")

survey18 %>%
  filter(!is.na(Country)) %>%
  count(Country) -> countries

highchart() %>%
hc_add_series_map(worldgeojson, countries, value = "n", joinBy = c('name','Country'))  %>%
  #hc_colors(c("darkorange", "darkgray")) %>%
    hc_colorAxis(stops = color_stops()) %>%
    hc_title(text = "Countries in World Map") %>%
    hc_subtitle(text = "This is beauty")
```



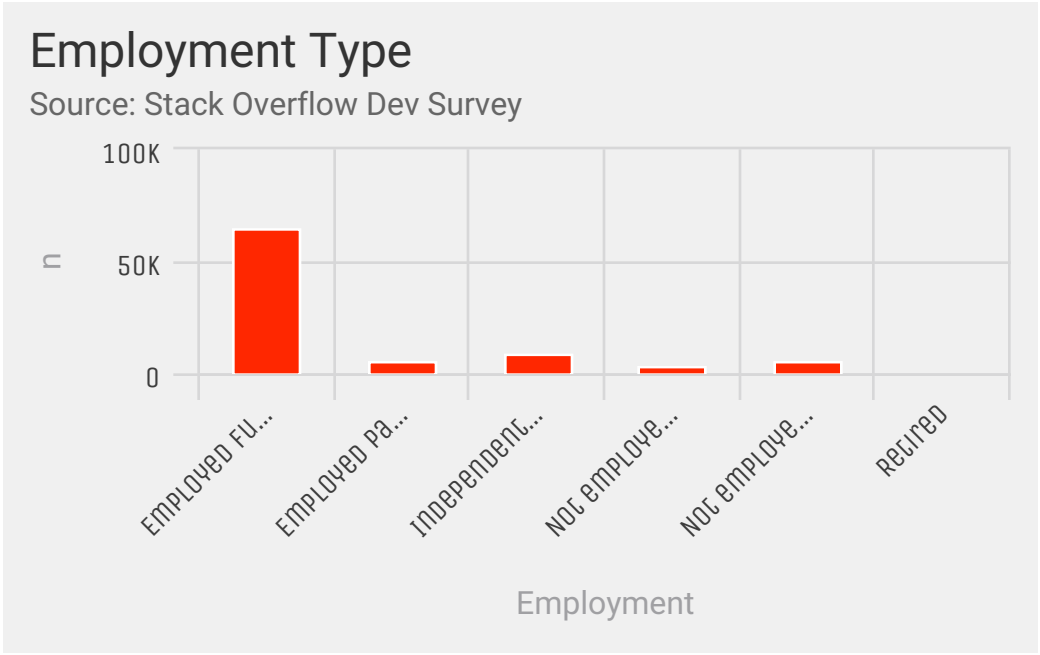
Themes

`Highcharter` also lets you aesthetically improve your charts with many inbuilt themes with just one extra line of code.

538

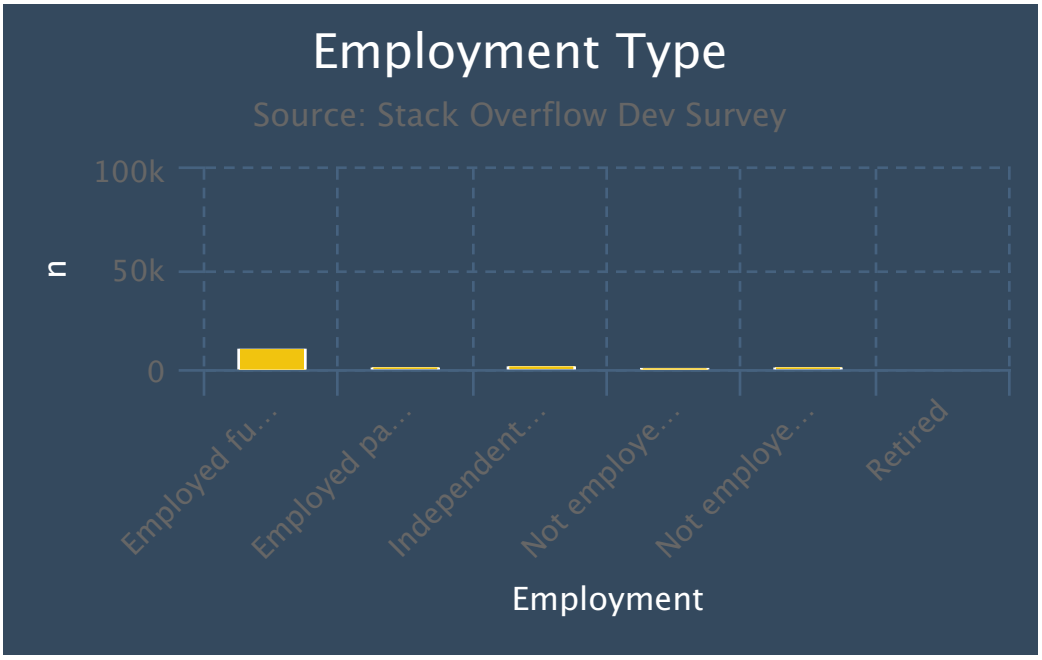
First the famous Five Thirty Eight Theme:

```
survey18 %>%
  filter(!is.na(Employment)) %>%
  count(Employment) %>%
  hchart('column',hcaes('Employment','n')) %>%
  hc_title(text = "Employment Type") %>%
  hc_subtitle(text = "Source: Stack Overflow Dev Survey") %>%
  hc_add_theme(hc_theme_538())
```



A dark theme

```
survey18 %>%
  filter(!is.na(Employment)) %>%
  count(Employment) %>%
  hchart('column',hcaes('Employment','n')) %>%
  hc_title(text = "Employment Type") %>%
  hc_subtitle(text = "Source: Stack Overflow Dev Survey") %>% hc_add_theme(hc_theme_flatdark())
```



Chalk Theme

```
survey18 %>%
  filter(!is.na(Employment)) %>%
  count(Employment) %>%
  hchart('column',hcaes('Employment','n')) %>%
  hc_title(text = "Employment Type") %>%
  hc_subtitle(text = "Source: Stack Overflow Dev Survey") %>%  hc_add_theme(hc_theme_chalk())
```