Principles of Data Visualization and Introduction to ggplot2

I have provided you with data about the 5,000 fastest growing companies in the US, as compiled by Inc. magazine. Please use the data available at
https://github.com/charleyferrari/CUNY_DATA608/tree/master/lecture1/Data

- This assignment must be done in a single R script with ggplot2. Use of dplyr or plyr is encouraged but not required.
- Images for question 'x' must be saved to the lecture1 directory as 'FigureX.jpg' or 'FigureX.png' (ie for Question 1, Figure1.jpg, etc).

Questions

1. Create a graph that shows the distribution of companies in the dataset by State (ie how many are in each state). There are a lot of States, so consider which axis you should use assuming I am using a 'portrait' oriented screen (ie taller than wide).
2. Let's dig in on the State with the 3rd most companies in the data set. Imagine you work for the state and are interested in how many people are employed by companies in different industries employ. Create a plot of average employment by industry for companies in this state (only use cases with full data (user R's complete.cases() function). Your graph should show how variable the ranges are, and exclude outliers.
3. Now imagine you work for an investor and want to see which industries generate the most revenue per employee. Create a chart makes this information clear.