# Review sheet

* Why is it important to learn data visualization?
  + The example of the historical cholera outbreak in London and its historical importance
  + Voronoi diagram: a simple way to improve the cholera map even further. What is it and how does it work?
    - Are you able to draw a Voronoi diagram yourself?
* History of visualizations
* Visualization integrity
  + Various types of misleading visualizations
  + How can a visualization lie?
* Perception and design
  + Psychophysics
  + Weber's law
  + Visual encodings
    - What is visual encoding?
  + Stevens' power law
    - What is Stevens' power law?
    - Why is Stevens' power law important?
    - What are the implications of Stevens' power law?
  + Colors
    - Why are they important?
    - How to encode colors? What are the systems and how do they work?
    - What are the basic concepts in colors?
    - Main usage types
    - How should we pick colors for categorical data?
    - How should we pick colors for quantitative data?
    - Accessibility of visualizations when using colors
  + Design principles
    - What are the gestalt principles? How can they be used and why does it matter?
    - Why are they important?
    - Data-ink ratio
    - Chart junk
    - Pre-attentive processing
* Data
  + Data types
    - Major data types and classification
  + Tidy data
    - What is tidy data?
    - Why is it important?
    - Can you make a dataset tidy?
* 1D (univariate) visualization
  + 1D scatter plot (strip chart)
  + Boxplot
    - Key ideas: Essential components
    - How to draw a boxplot?
    - What are the pros and cons of box plots?
  + Histogram
    - Key ideas: Basic logic of what they represent
    - How to draw a correct histogram?
    - Choosing bins
    - Cumulative histogram
  + CDF & CCDF
    - Key ideas and how they work
    - How to draw one correctly given a small set of data
  + Density Estimation
    - Why and how
    - Parametric vs. non-parametric approaches
    - Kernel density estimation
      * Choices in KDE
  + Characteristics and comparison of univariate density visualization methods
  + Interpolation
  + Extrapolation
  + Moving average
  + Regression
* Log-scale
  + Log-scale basics
  + Why is it important?
  + Visualizing ratios
  + Drawing a histogram in a log-scale
  + CDF, CCDF
* 2D (bivariate) visualizations
  + Scatterplot
  + Histogram (heatmap)
* High dimensional data
  + Scatterplot matrix
  + Parallel coordinates
  + Radar charts
  + Dimensionality reduction
    - Data manifold. Why can dimensionality reduction work?
    - Common techniques: how do they work? What kinds of information do they use?
    - How do PCA, MDS, t-SNE, UMAP work?
    - What are their key properties and how should we use them?
* Maps
  + Map projection.
  + Mercator projection
  + Why do we need a cartogram? When?
  + What is the basic idea of a cartogram? How does it work?
* Texts
  + bag-of-the-words model
  + Word cloud
* Networks
  + Node-link diagram and force-directed layout algorithm
  + Trees and treemap