



Data Visualization with R



Matthew Renze
Data Science Consultant
Renze Consulting

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mobilow et se rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

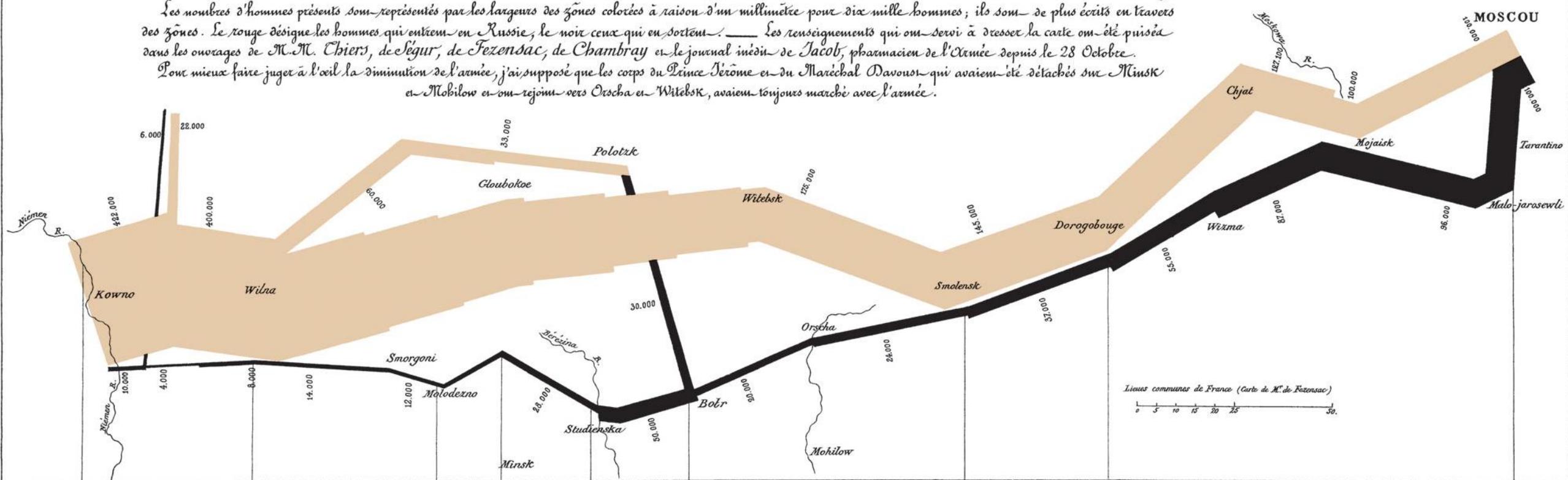
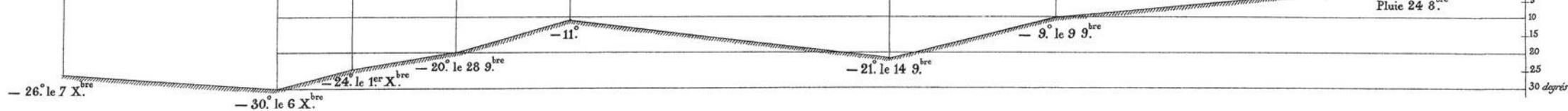


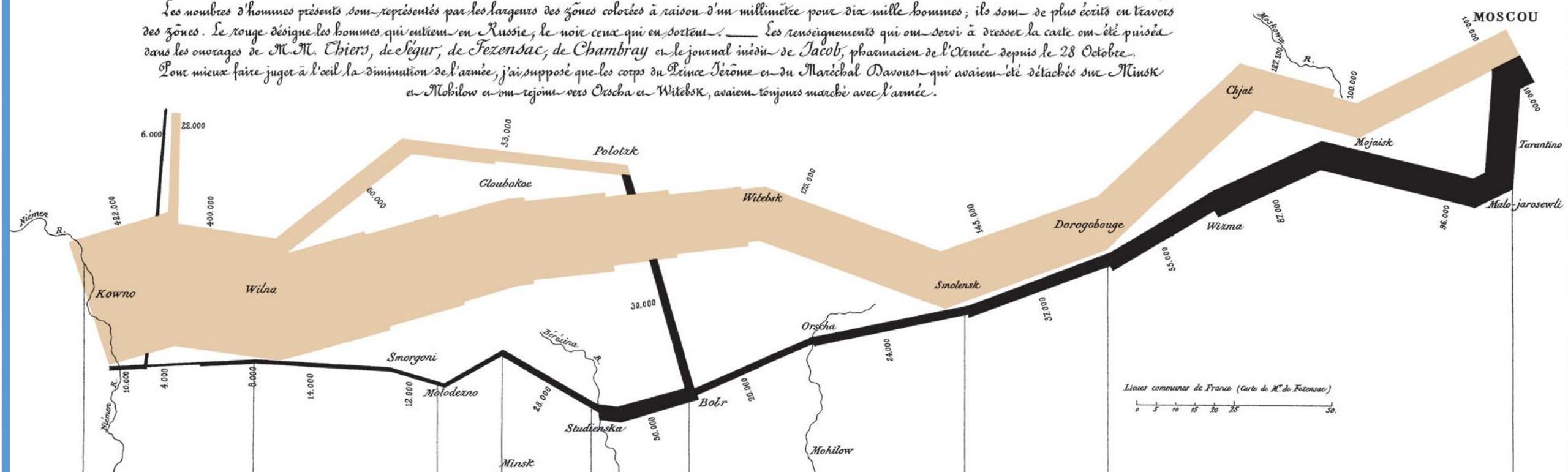
TABLEAU CRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussees en retraite. Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et qui rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.



Lieux communs de France (Carte de M. de Fezensac)
 1 5 10 15 20 25 50.

TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



Les cosaques passent au galop
le Niémen gelé.

Avec. par Regnier, 8 Rue St Marie St Gm à Paris.

Imp. Lith. Regnier et Bourdet.

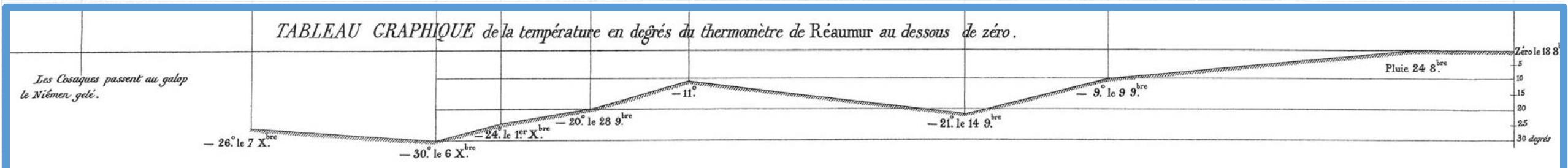
Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dessiné par M. Minard, Ingénieur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes, ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chier, de Léger, de Fézenda, de Chambray et le journal médical de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et avaient rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.



Lignes communes de France (Carte de M. de Fonsseca)



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mogilow et se rejoignaient vers Orscha et Witebsk, avaient toujours marché avec l'armée.

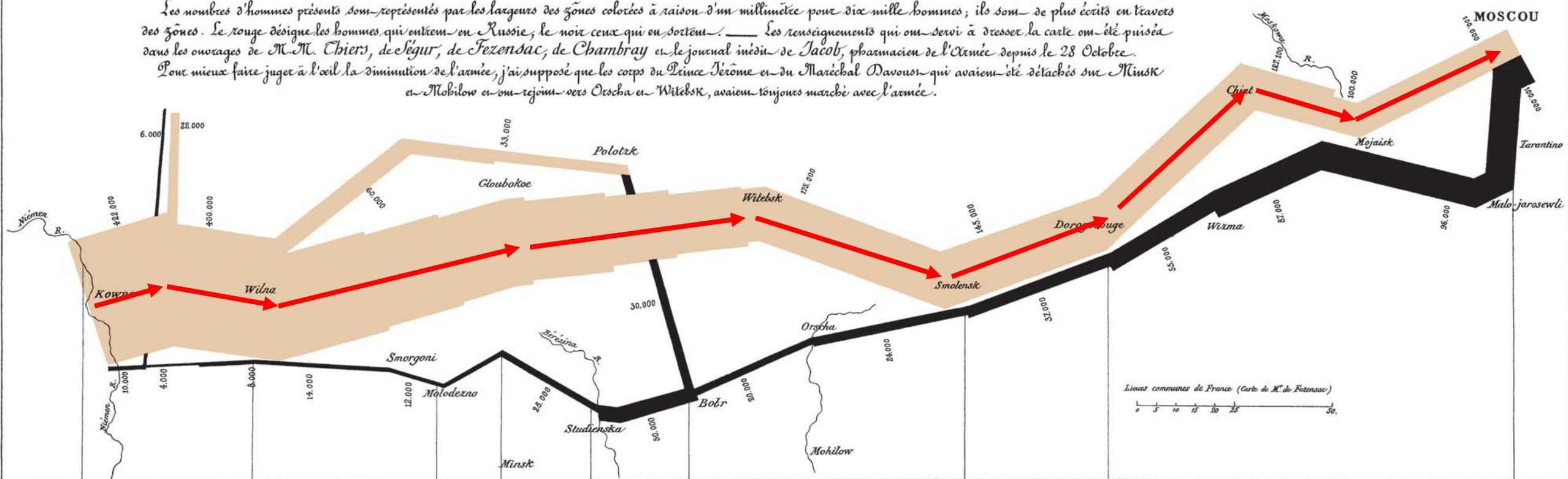


TABLEAU CRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les cosaques passent au galop
le Niemen gelé.

— 26° le 7 X^{bre}.
 — 30° le 6 X^{bre}.
 — 24° le 1^{er} X^{bre}.
 — 20° le 28 9^{bre}.
 — 11°.

— 21° le 14 9^{bre}.
 — 8° le 9 9^{bre}.

Zéro le 18 8^{bre}.
 Pluie 24 8^{bre}.
 5
 10
 15
 20
 25
 30 degrés

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mobilow et se rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

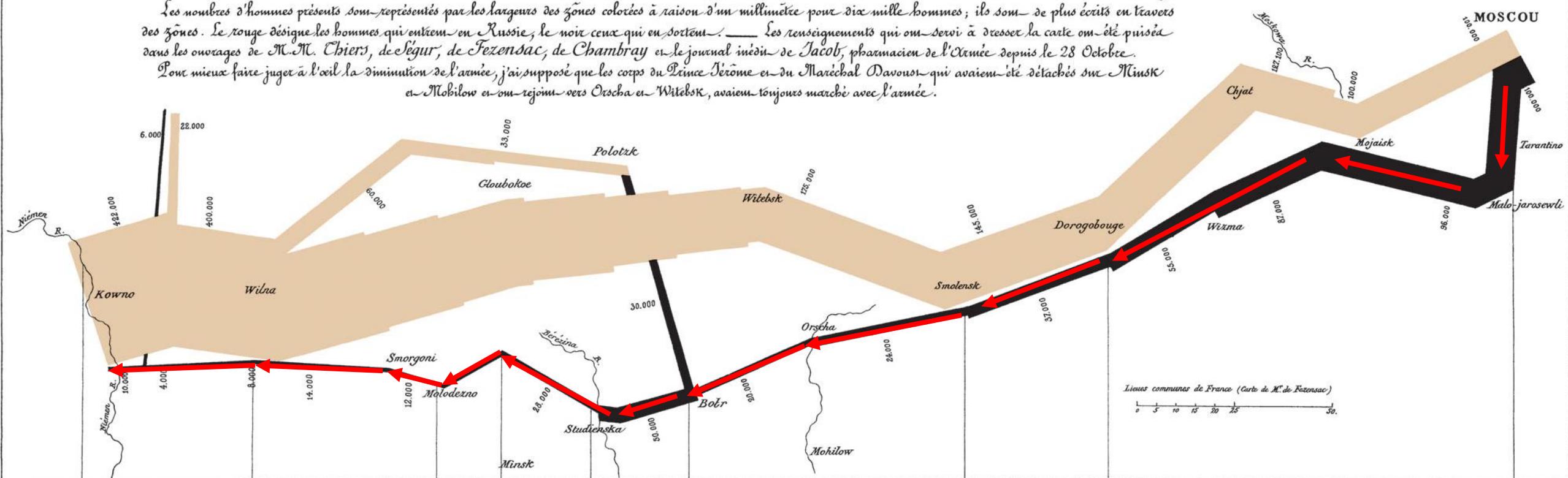


TABLEAU CRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les cosaques passent au galop
le Niemen gelé.

— 26° le 7 X^{bre}.
 — 30° le 6 X^{bre}.
 — 24° le 1^{er} X^{bre}.
 — 20° le 28 9^{bre}.
 — 11°.

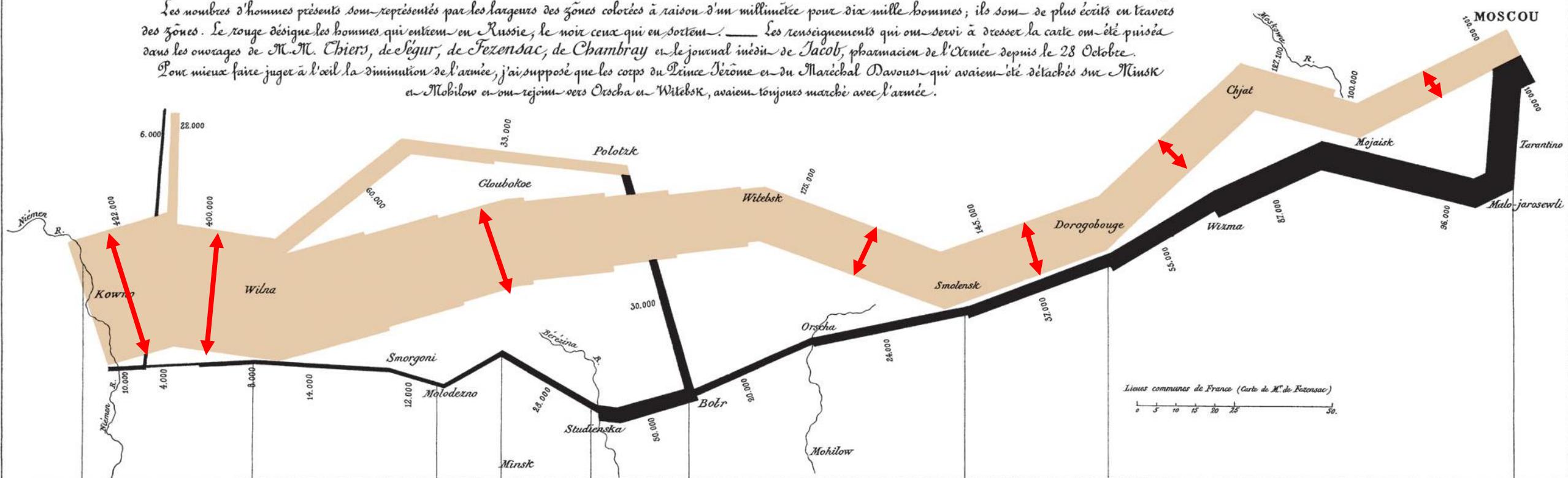
— 21° le 14 9^{bre}.
 — 8° le 9 9^{bre}.

Zéro le 18 8^{bre}.
 Pluie 24 8^{bre}.
 5
 10
 15
 20
 25
 30 degrés

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

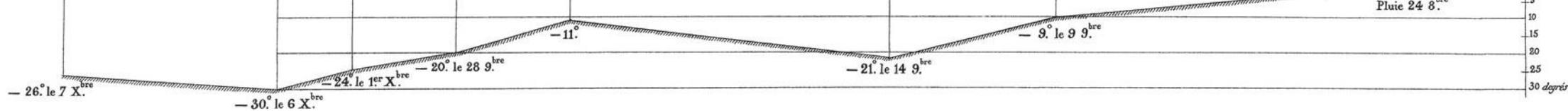
Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mogilow et se rejoignaient vers Orscha et Witebsk, avaient toujours marché avec l'armée.



Lieux communs de France (Carte de M. de Fezensac)

TABLEAU CRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

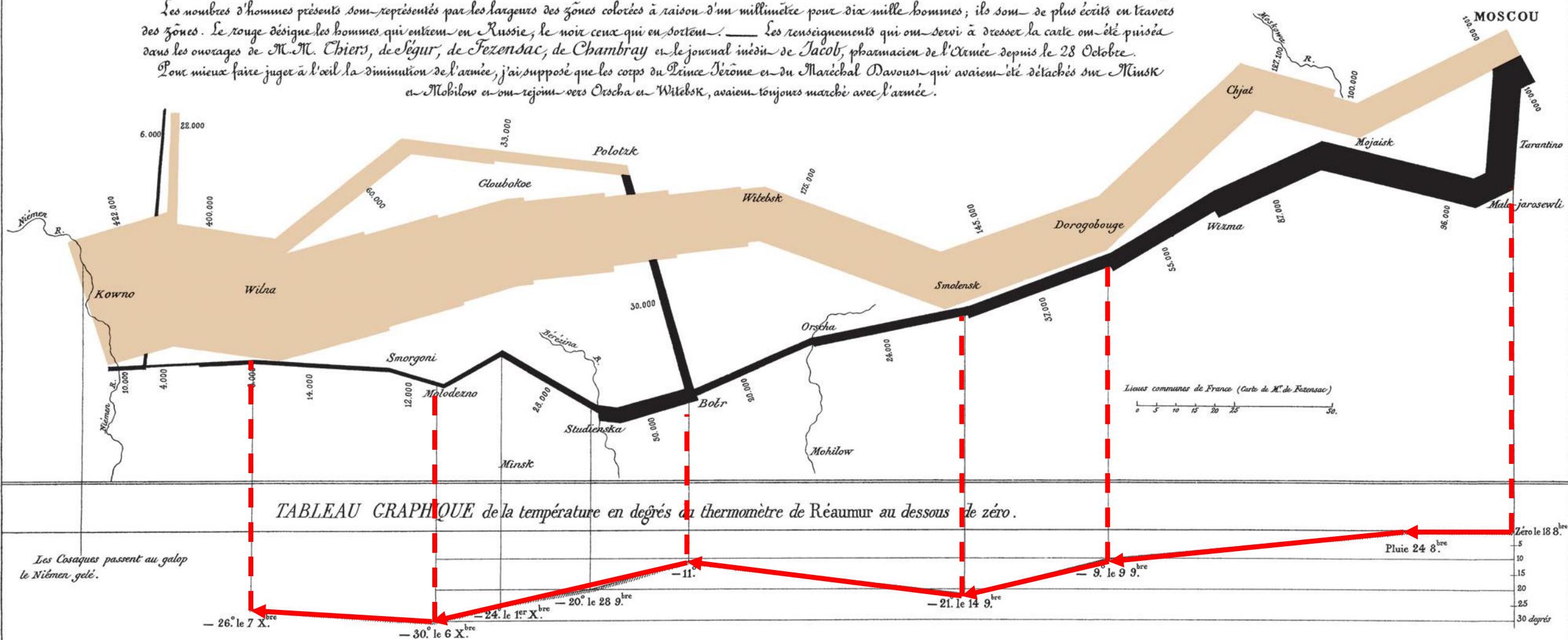


Les cosaques passent au galop
le Niemen gelé.

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mogilow et qui rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dessinée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite
Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui ont été en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à desser la carte ont été puisés dans les ouvrages de M. M. Chier, de Séguir, de Fézensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk en Mobilow et se rejoignaient vers Orscha et Wilejka, avaient toujours marché avec l'armée.

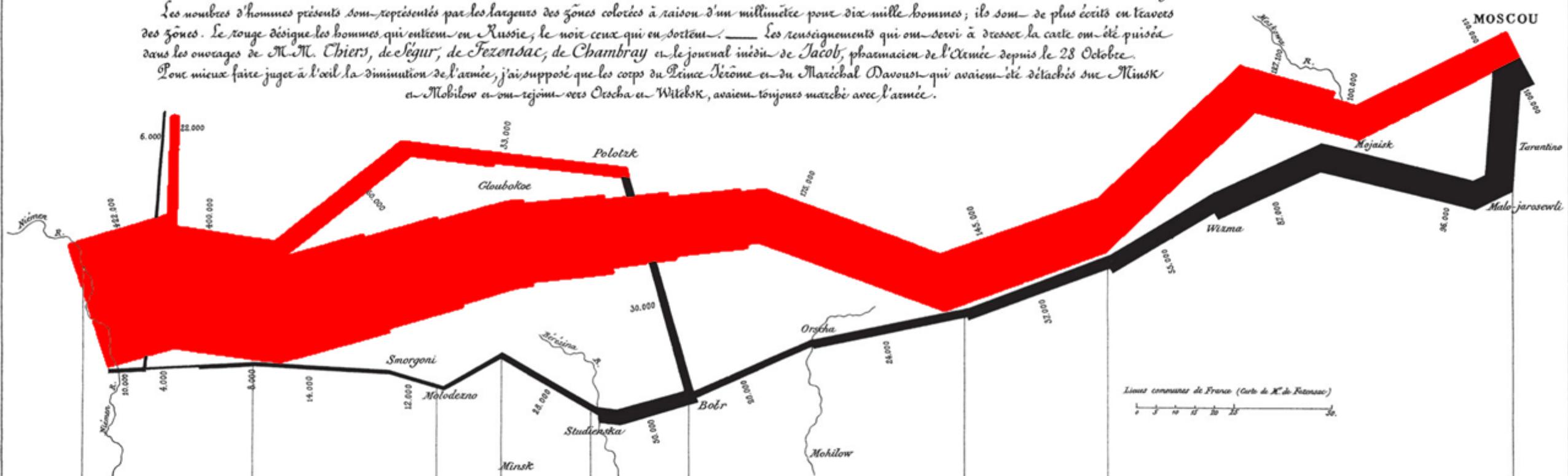


TABLEAU CRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop
le Niemen gelé.

- 26° le 7 X.^{bre}

- 30° le 6 X.^{bre}

- 24° le 1^{er} X.^{bre}

- 11°

- 21° le 14. 9.^{bre}

- 9° le 9. 9.^{bre}

Pluie 24. 8.^{bre}

Zéro le 18. 8.^{bre}

5

10

15

20

25

30 degrés

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dessinée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite
Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui ont été en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Léger, de Fézenda, de Chambray et le journal médical de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mogilow et qui rejoignirent Orla et Wilejka, avaient toujours marché avec l'armée.

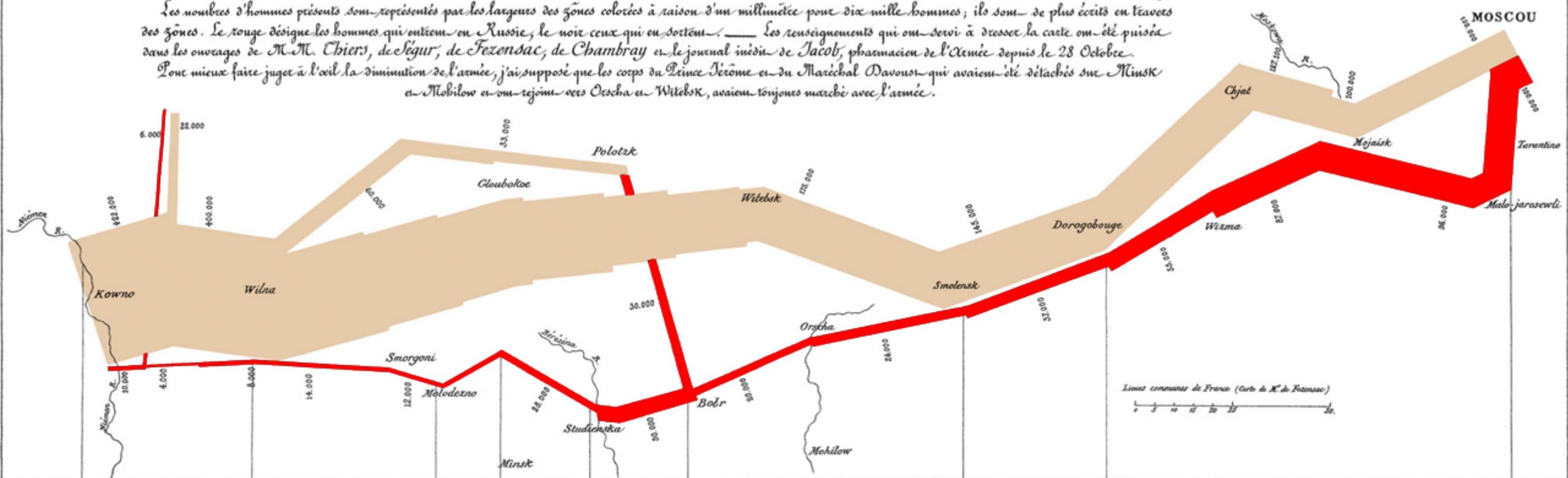
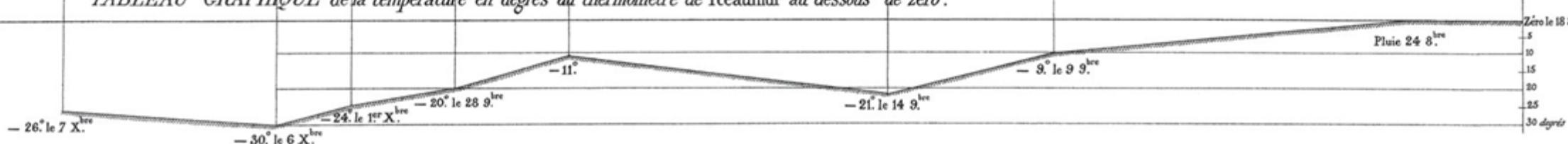


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les cosaques passent au galop
le Niemen gelé.



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812 - 1813.
 Dessinée par M. Minard, Ingénieur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes, ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chicher, de Léglise, de Fezensac, de Chambray et le journal médical de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Moscou au Mohilow et avaient rejoint Orscha et Witebsk, avaient toujours marché avec l'armée.

Rule #1: Never march on Moscow

Rule #2: Never fight a land war in Asia

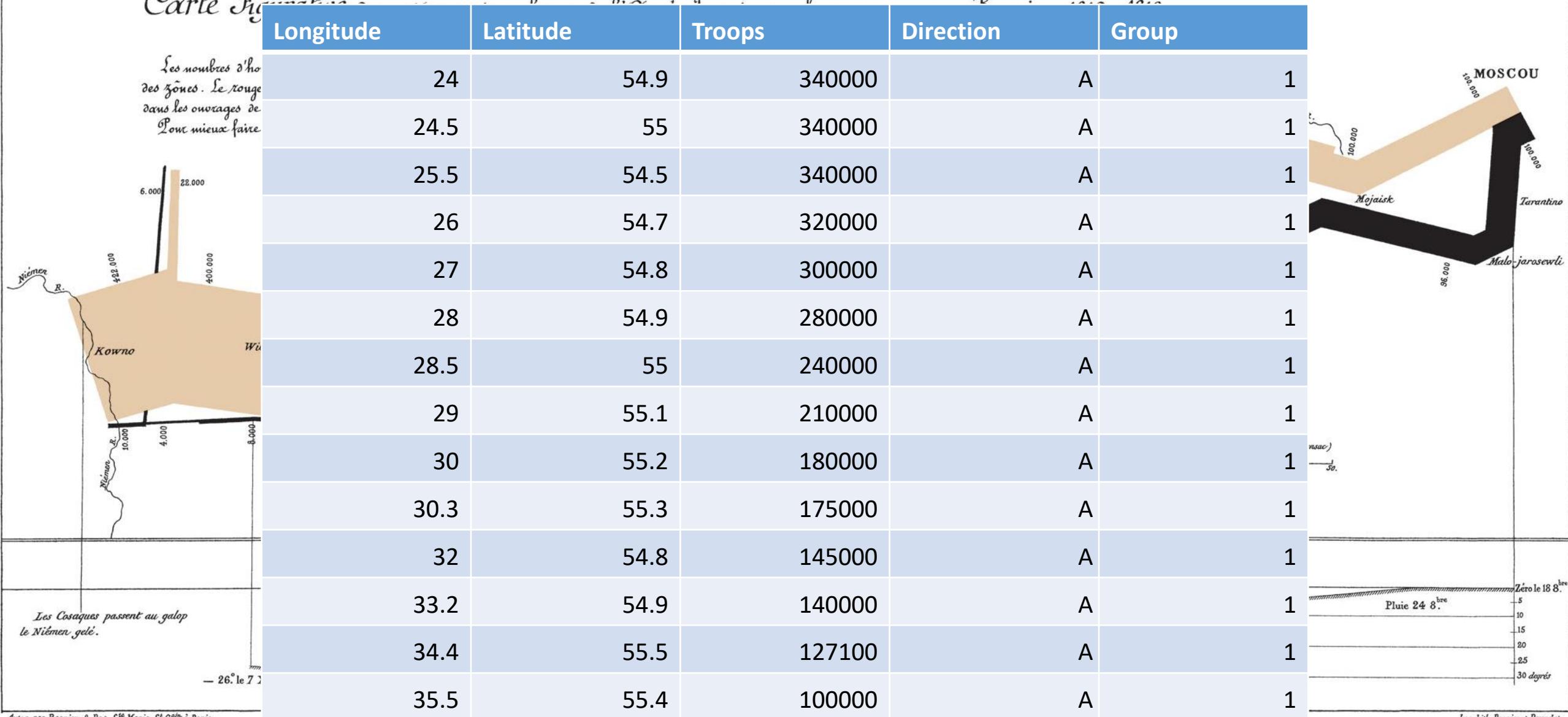
There is no victory... only a degree of loss.

TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les cosaques passent au galop le Niemen gelé.



Carte Figurative



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite
 Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mobilow et se rejoignaient vers Orscha et Witebsk, avaient toujours marché avec l'armée.

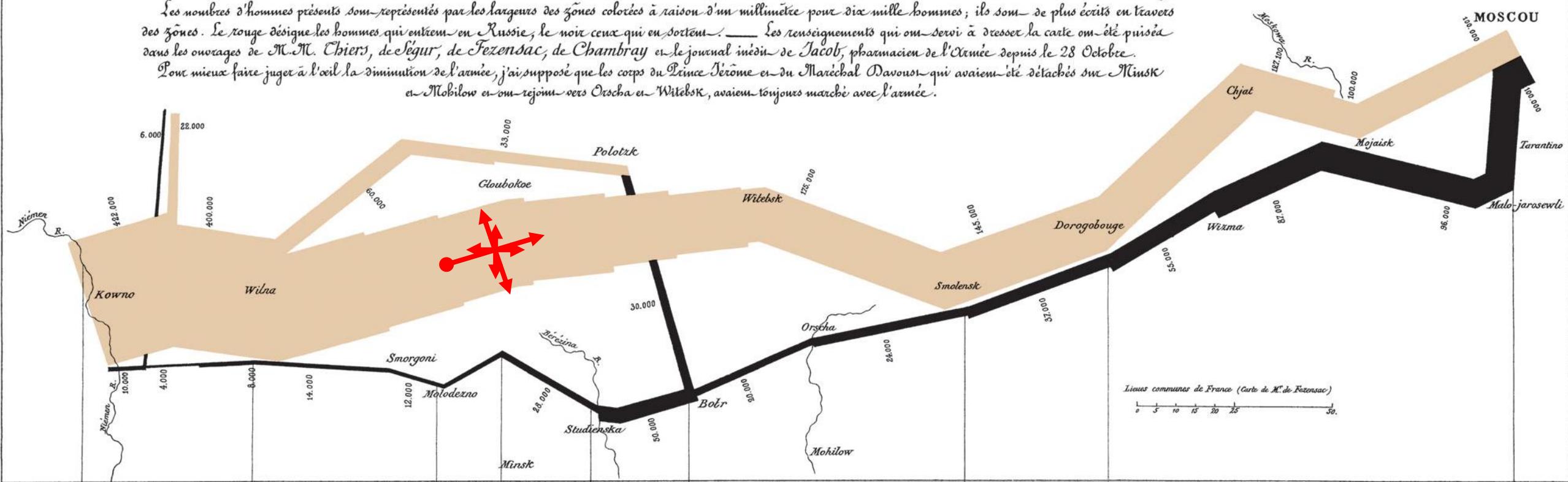
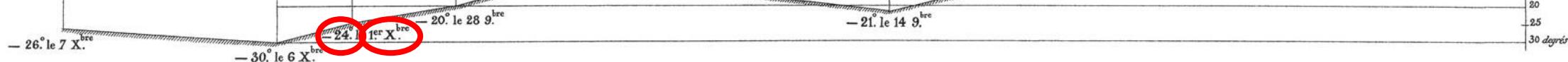


TABLEAU CRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les cosaques passent au galop
le Niemen gelé.



Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.
 Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite à Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mobilow et se rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

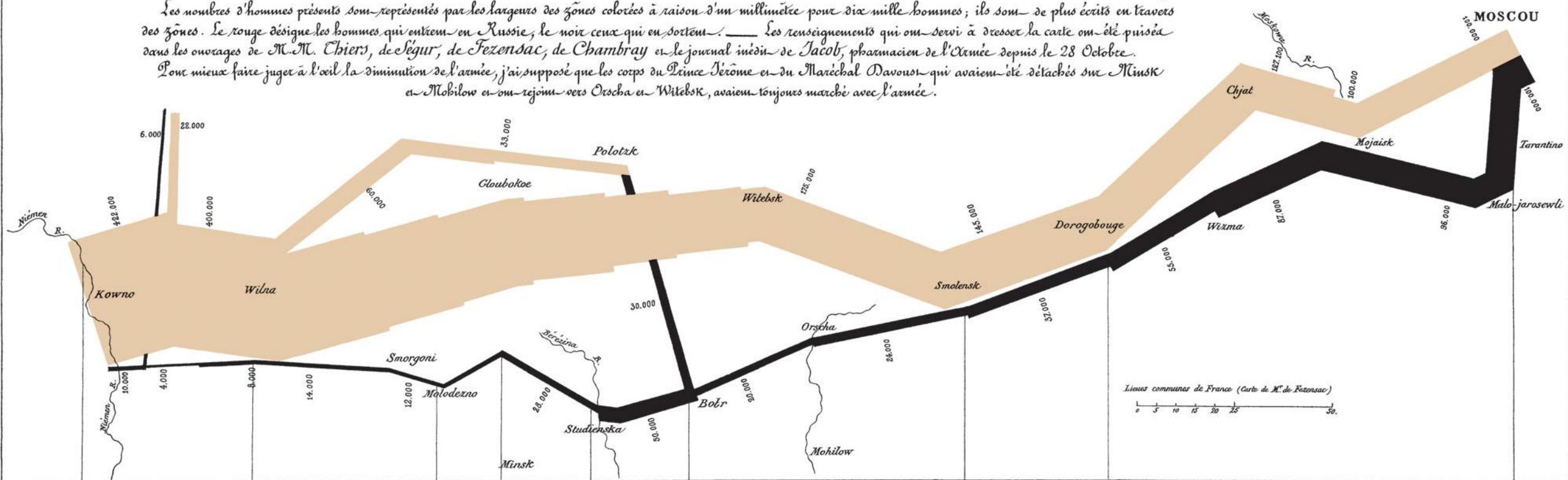
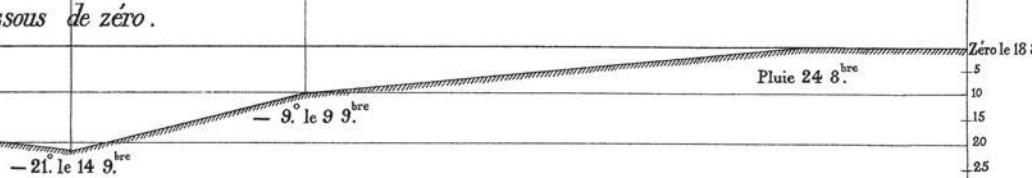


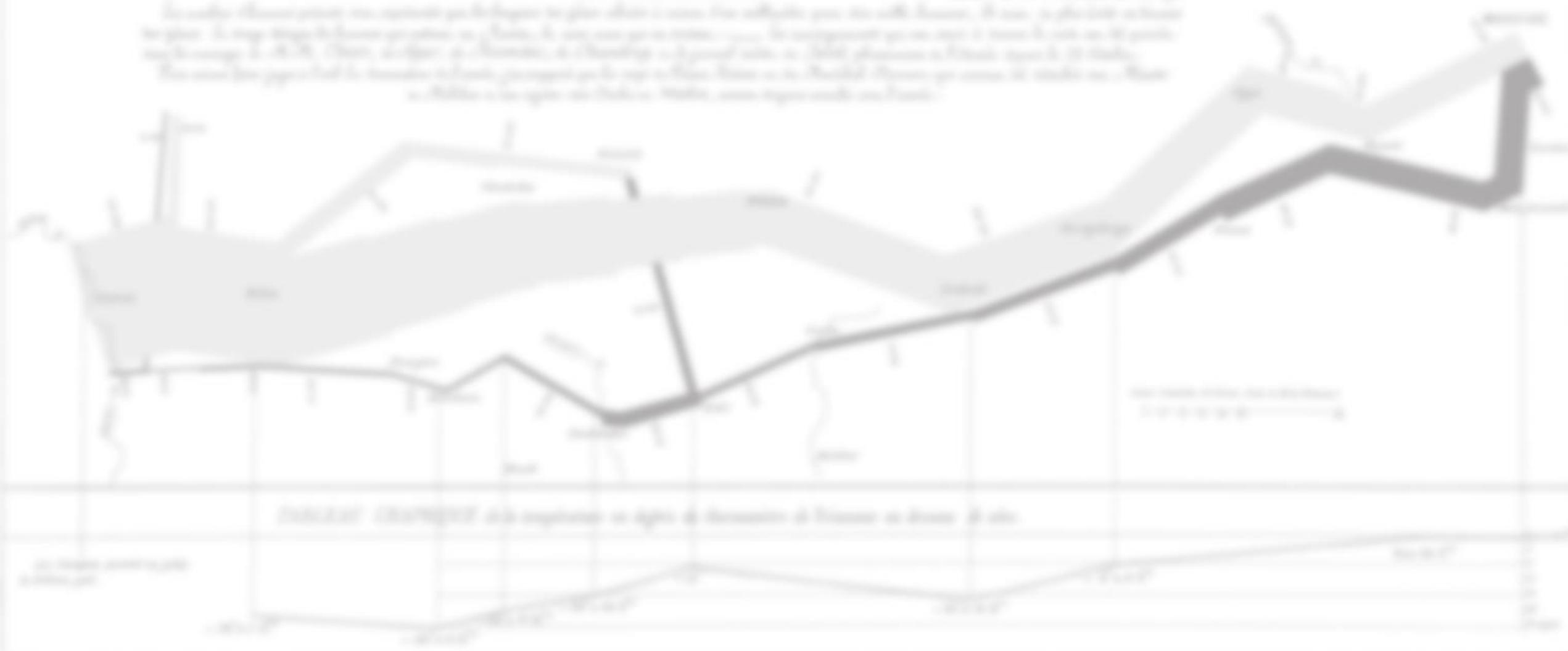
TABLEAU CRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

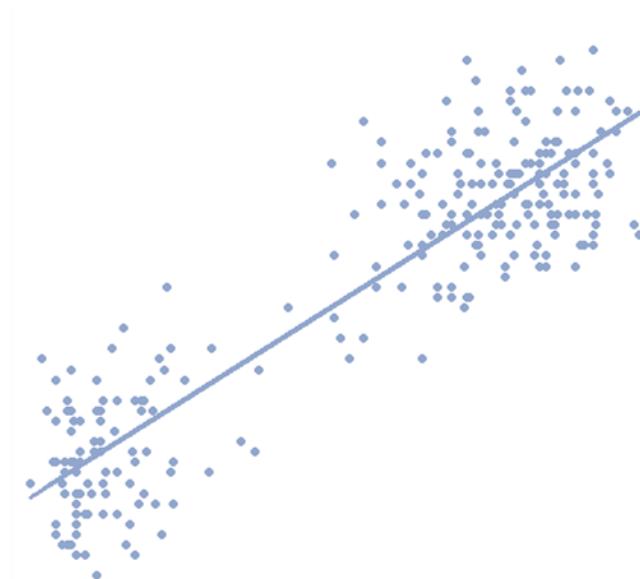
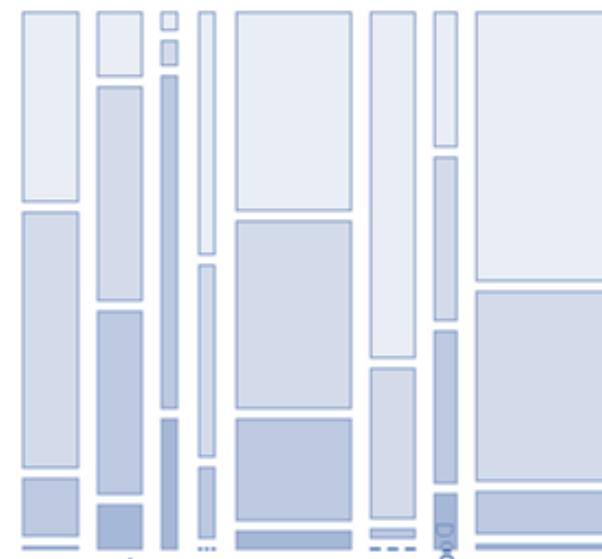
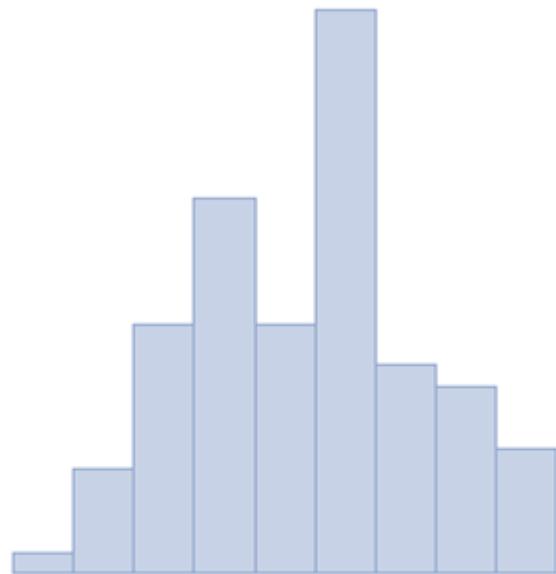
Les cosaques passent au galop
le Niemen gelé.

— 26° le 7 X^{bre}.
 — 30° le 6 X^{bre}.
 — 24° le 1^{er} X^{bre}.
 — 20° le 28 9^{bre}.
 — 11°.

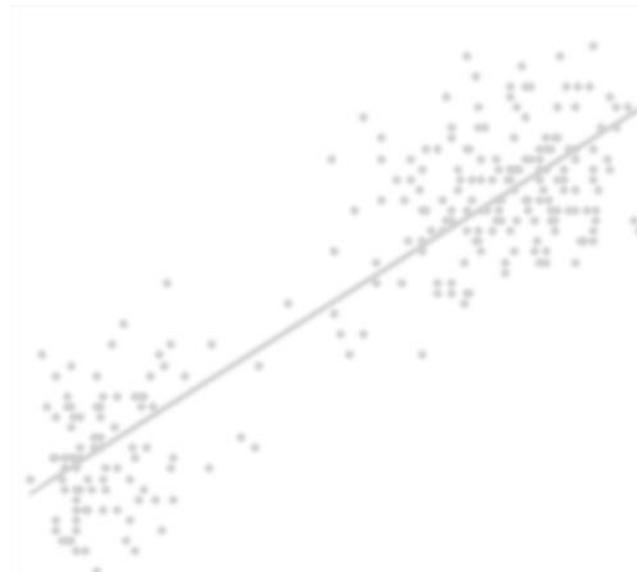


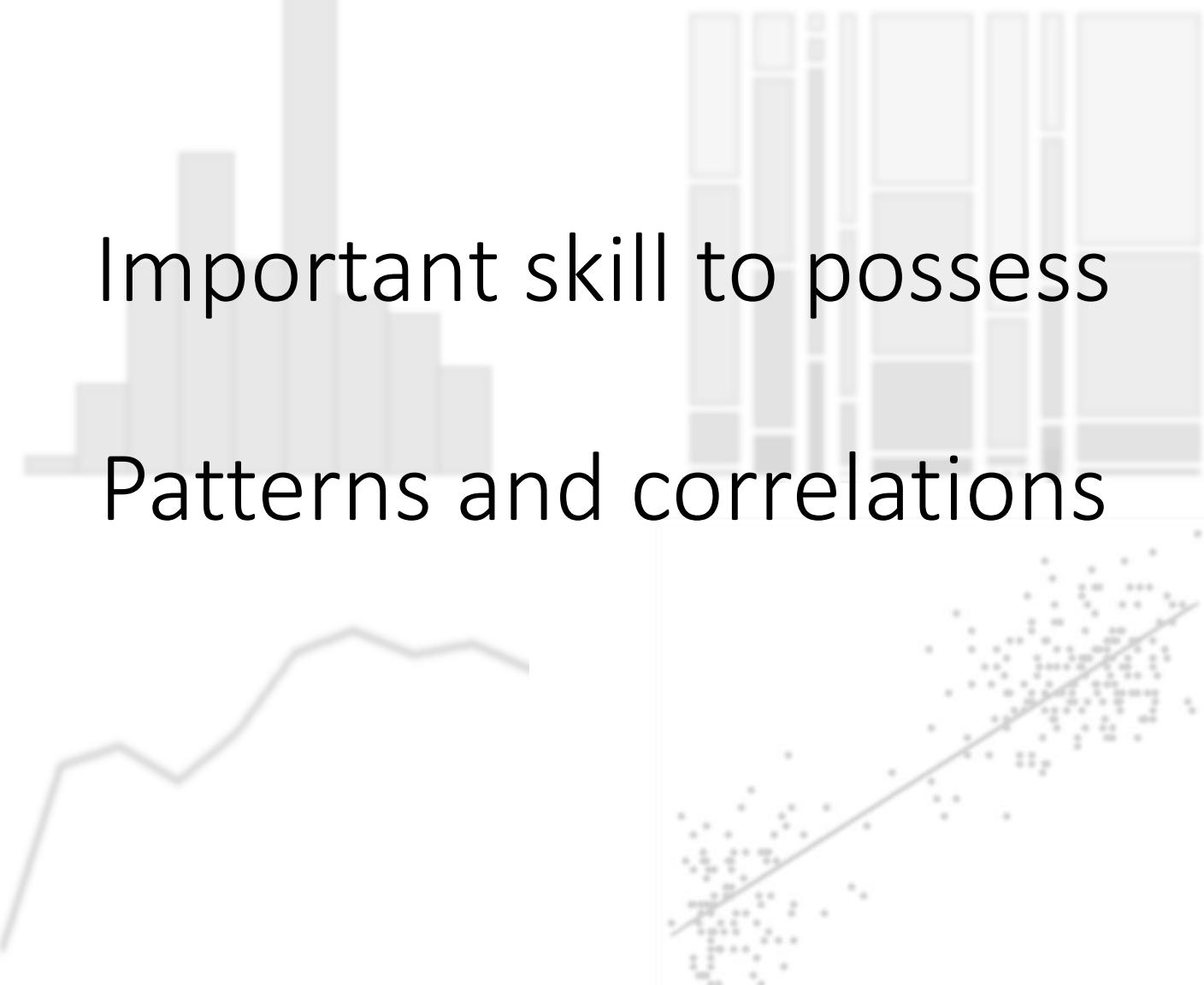
Costs of public services in general increase with time. This chart shows how the cost of running a local government has increased over time. It is likely that the budget for public services will continue to grow in the future. This is because there are more people in the country, and they require more services.





Important skill to possess





Important skill to possess

Patterns and correlations



Important skill to possess

Patterns and correlations

Tell a story with data

Overview

Introduction to R

Intro to Data Visualization

Types of Data Visualizations

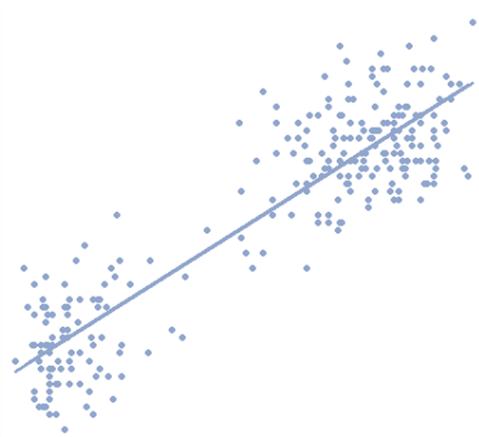
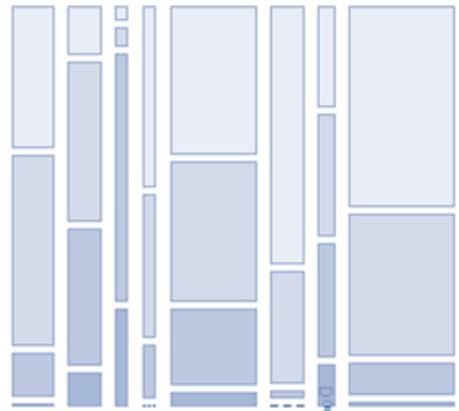
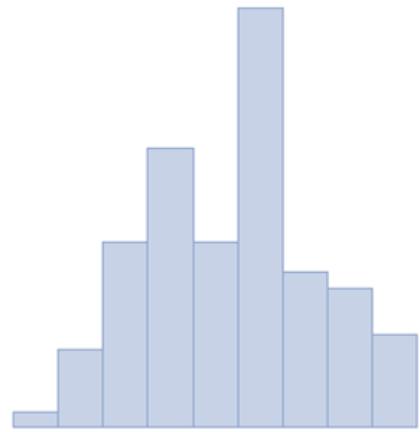
- Visualizing One Variable

- Visualizing Two Variables

- Visualizing Many Variables

- Visualizing Other Types of Data

Beyond the Basics



Introduction to R

What is R?

Open source

Language and environment

Numerical and graphical analysis

Cross platform



What is R?

- Active development
- Large user community
- Modular and extensible
- 10,000+ extensions



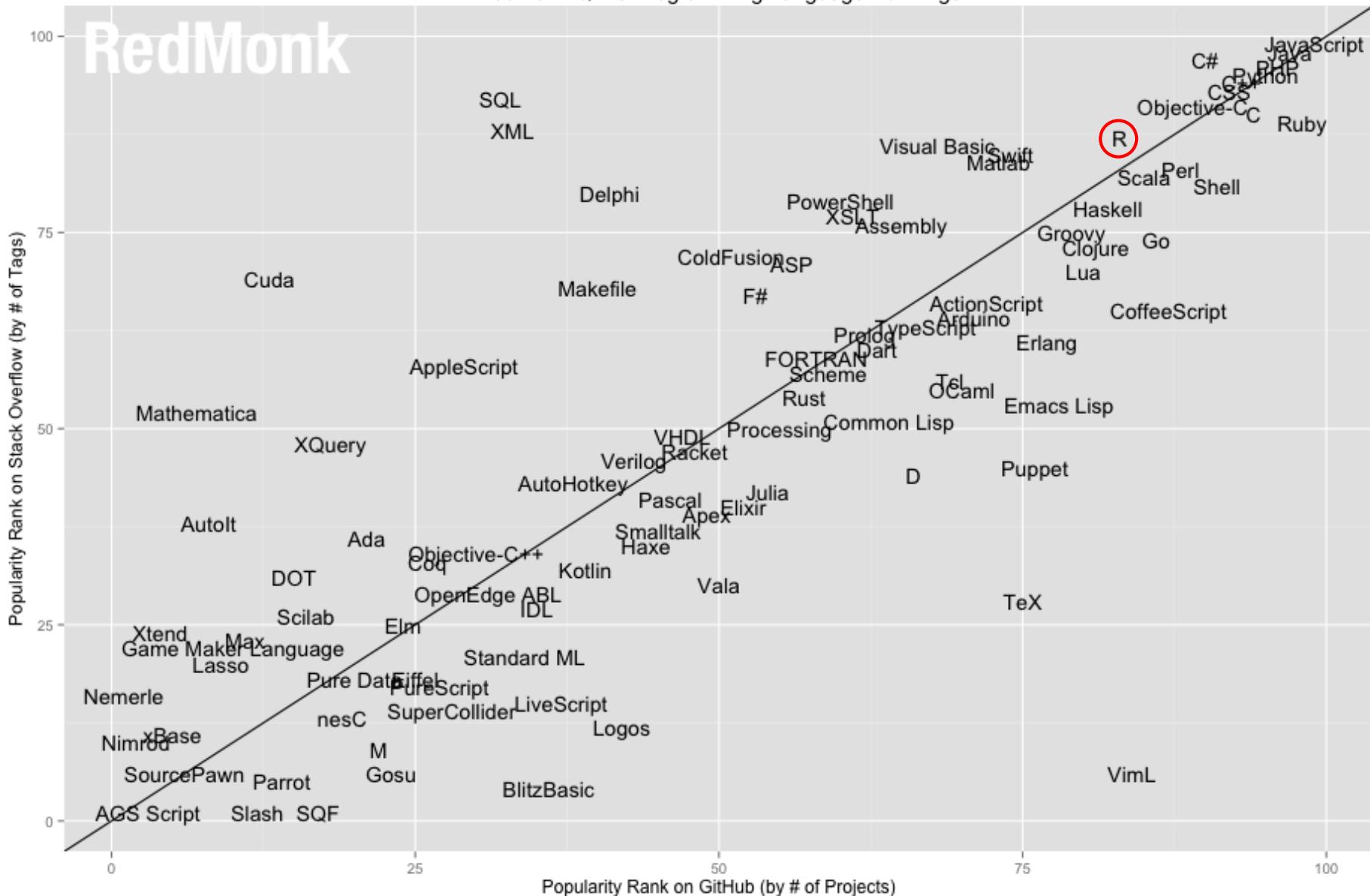
FREE



A low-angle photograph of the Statue of Liberty against a clear blue sky. Her right arm is raised high, holding a torch aloft. Her left arm is bent, holding a tablet or smartphone that displays the word "FREE".

FREE

RedMonk Q116 Programming Language Rankings



Source: <http://redmonk.com/sogrady/2016/07/20/language-rankings-6-16/>

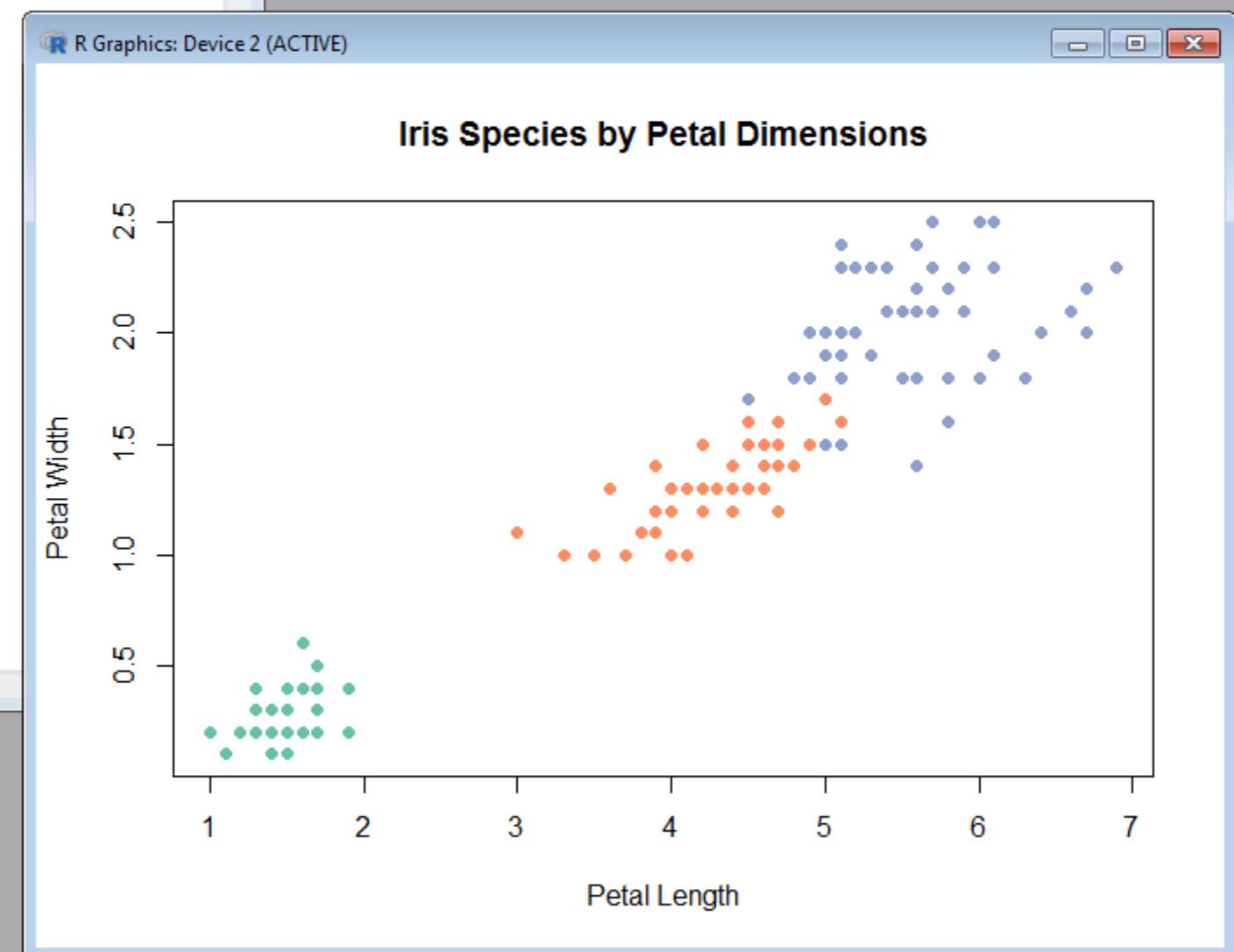


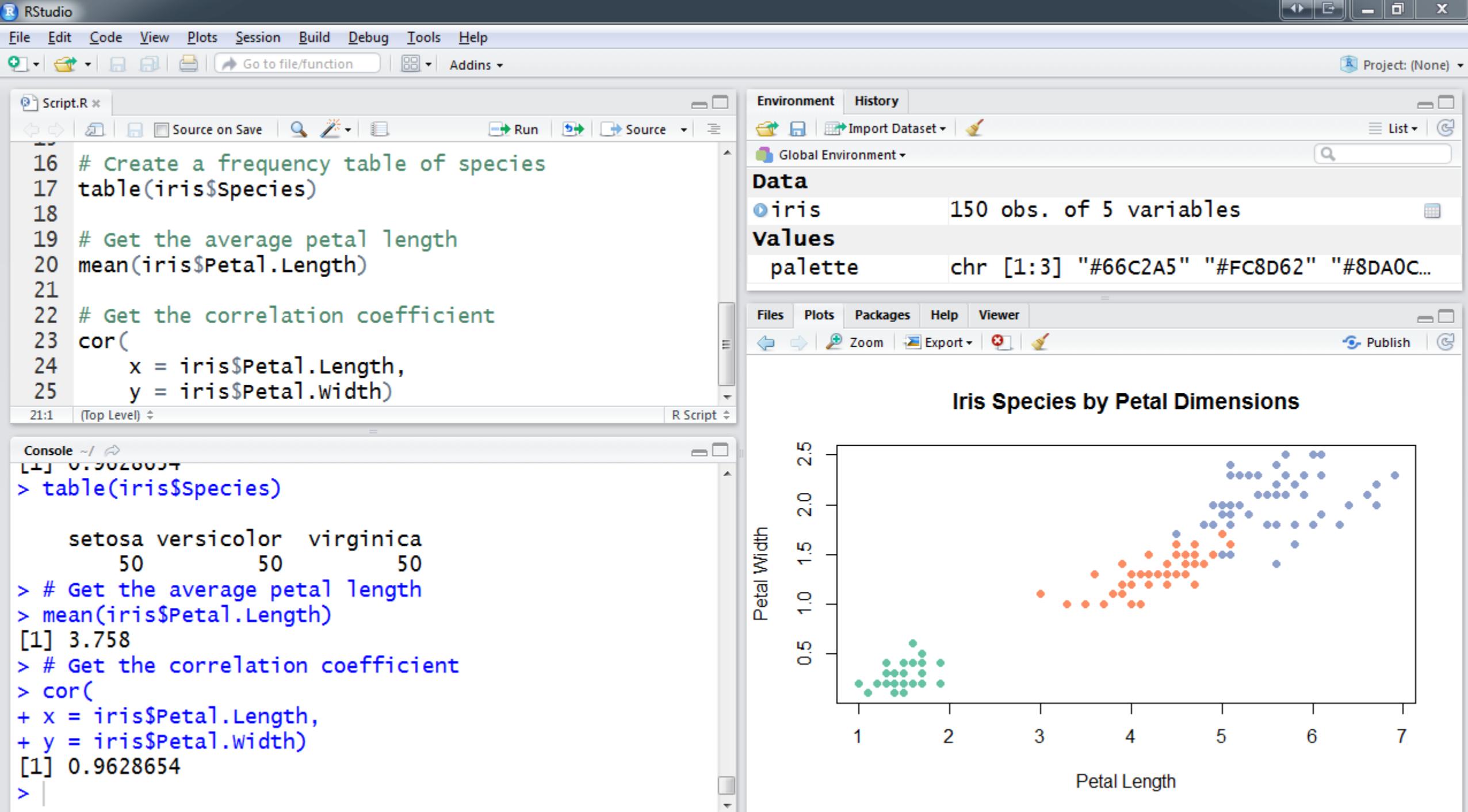
R Console

```
> # Create a plot of species by dimension
> plot(
+   x = iris$Petal.Length,
+   y = iris$Petal.Width,
+   pch = 19,
+   col = palette[as.numeric(iris$Species)],
+   main = "Iris Species by Petal Dimensions",
+   xlab = "Petal Length",
+   ylab = "Petal Width")
>
> # Create a frequency table of species
> table(iris$Species)

  setosa versicolor virginica 
      50       50       50 

>
> # Get the average petal length
> mean(iris$Petal.Length)
[1] 3.758
>
> # Get the correlation coefficient
> cor(
+   x = iris$Petal.Length,
+   y = iris$Petal.Width)
[1] 0.9628654
```





Script.R - Microsoft Visual Studio

File Edit View NCrunch Project Debug Team Tools Architecture Test ReSharper R Tools Analyze Window Help

Matthew Renze

Quick Launch (Ctrl+Q)

Script.R

```
main = "Iris Species by Petal Dimensions",
xlab = "Petal Length",
ylab = "Petal Width")

# Create a frequency table of species
table(iris$Species)

# Get the average petal length
mean(iris$Petal.Length)

# Get the correlation coefficient
cor(
  x = iris$Petal.Length,
  y = iris$Petal.Width)
```

R Interactive

```
> # Create a frequency table of species
> table(iris$Species)

  setosa versicolor virginica
      50          50         50
> # Get the average petal length
> mean(iris$Petal.Length)
[1] 3.758
> # Get the correlation coefficient
> cor(
+   x = iris$Petal.Length,
+   y = iris$Petal.Width)
[1] 0.9628654
>
```

Variable Explorer

Name	Value	Class	Type
iris	150 obs. of 5 variables	data.frame	list
palette	chr [1:3] "#66C2A5" "#FC8D62" "#8DA0CE"	character	character

R Plot

Iris Species by Petal Dimensions

Petal Width

Petal Length

Solution Explorer R Plot R Package Manager R Help

Error List Output Azure App Service Activity

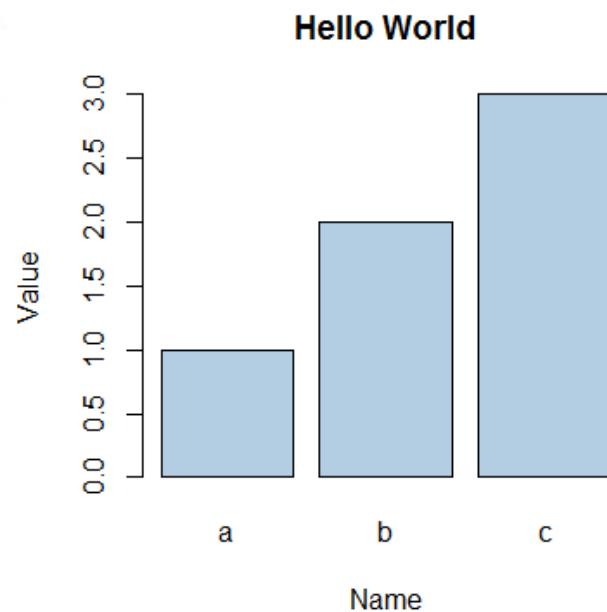
Ln 30 Col1 Ch1 INS ↑ 7 ↗ 0 ⌘ Root ⌘ master

Code Demo

Intro to Data Visualization

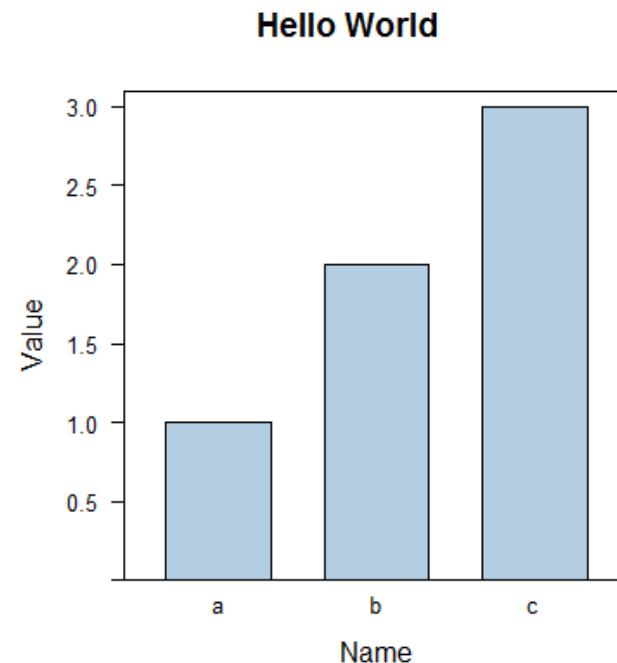
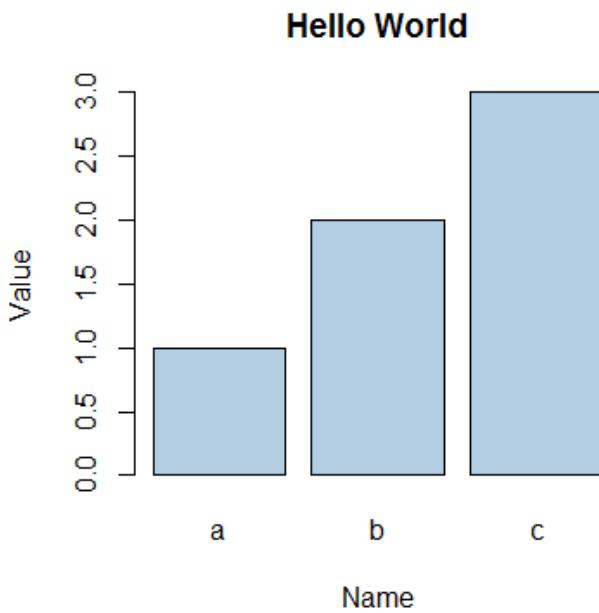
Plotting Systems in R

Plotting Systems in R



Base

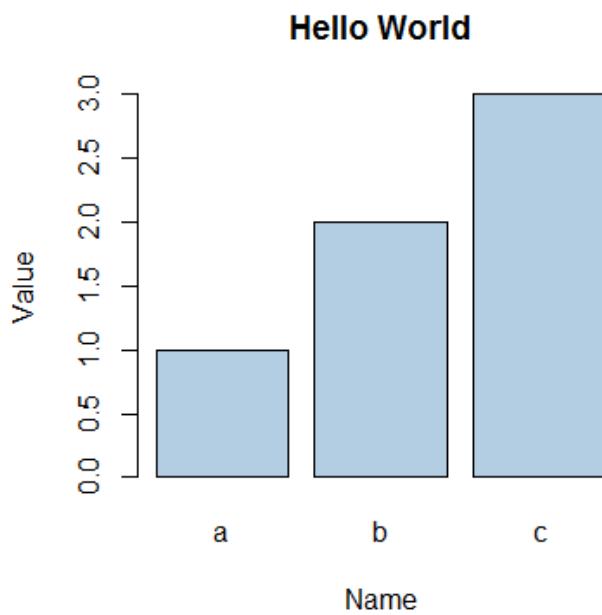
Plotting Systems in R



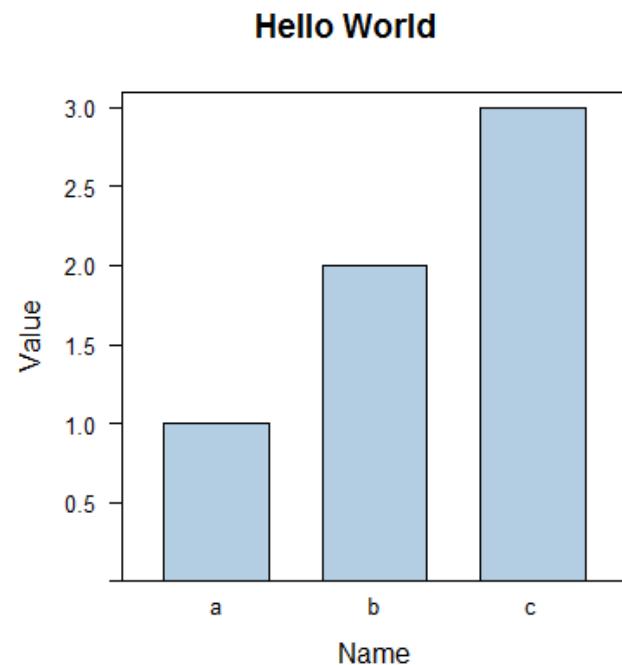
Base

Lattice

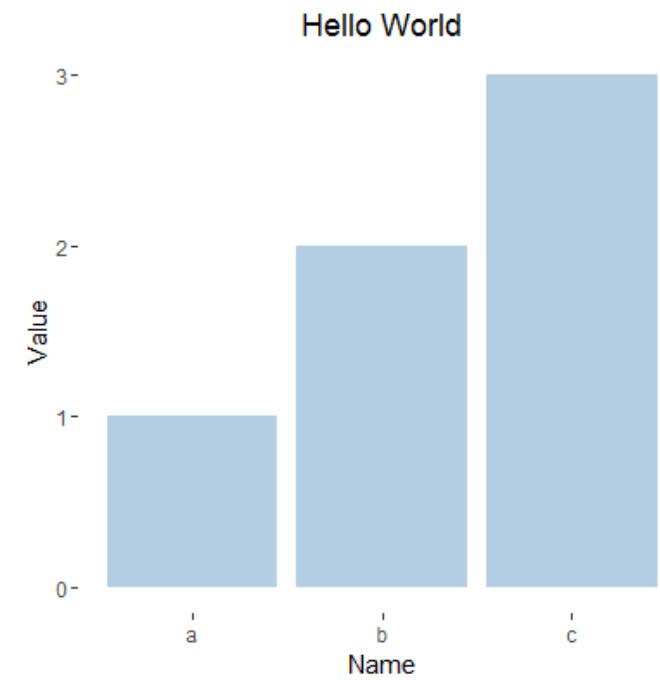
Plotting Systems in R



Base

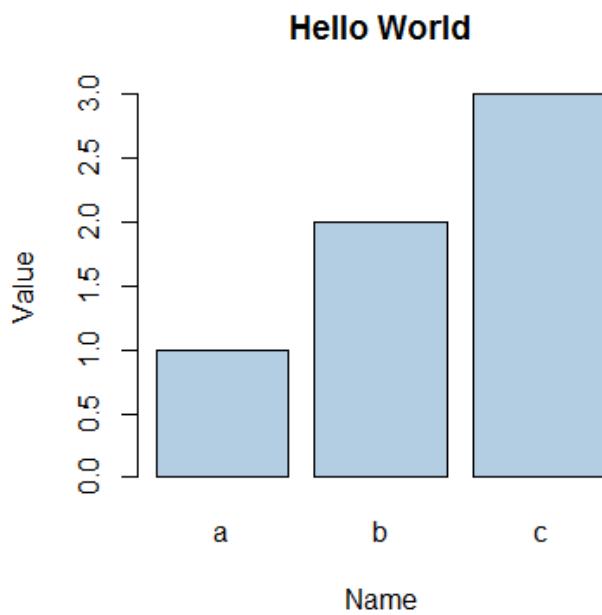


Lattice

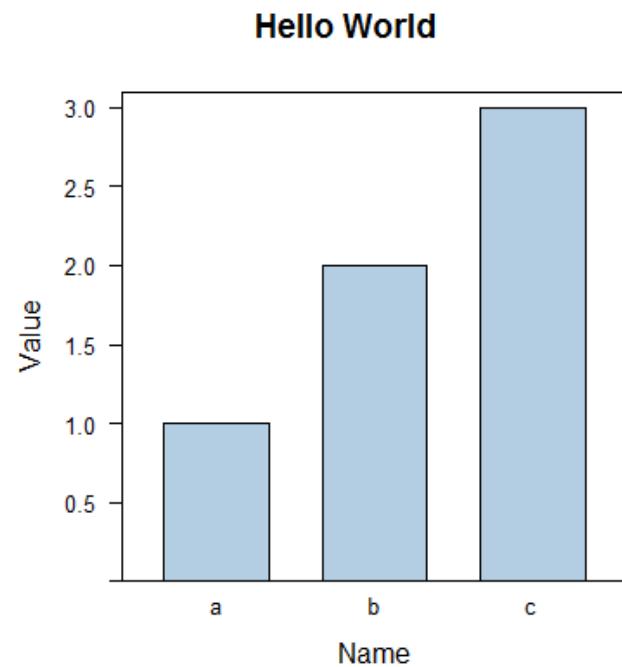


ggplot2

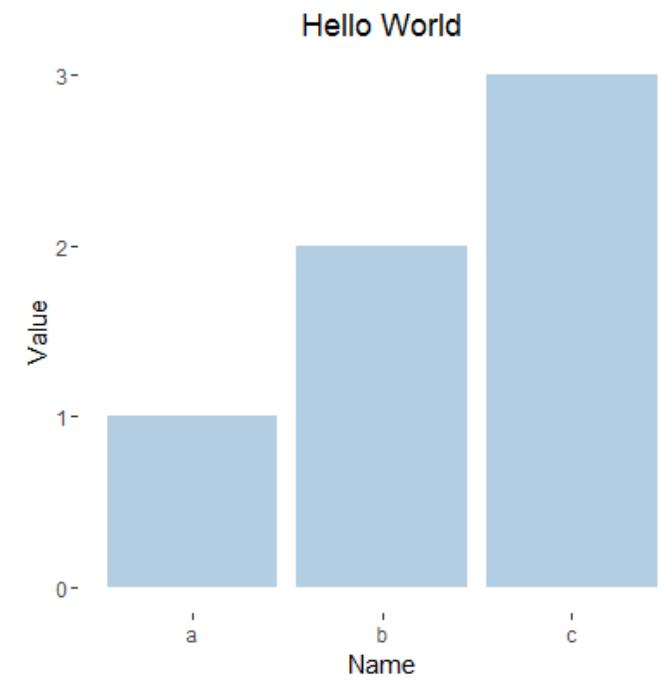
Plotting Systems in R



Base



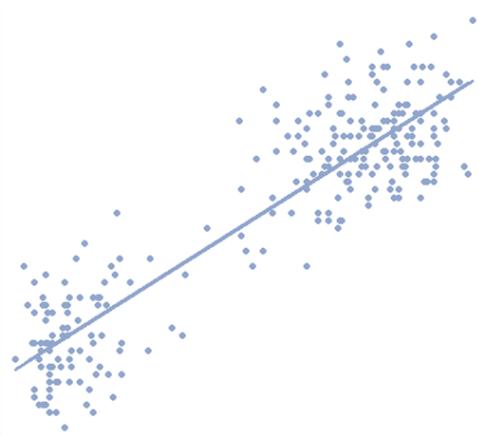
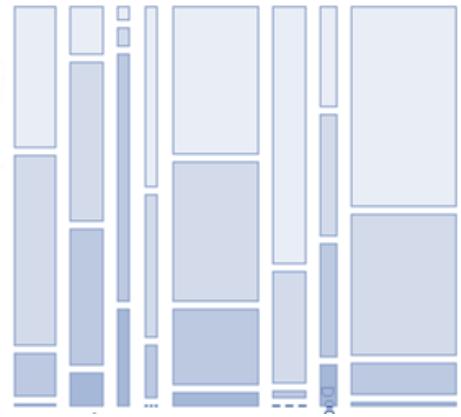
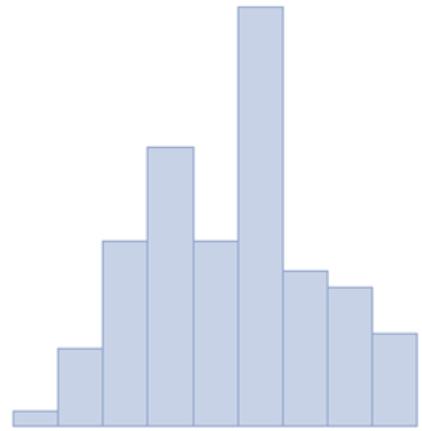
Lattice



ggplot2

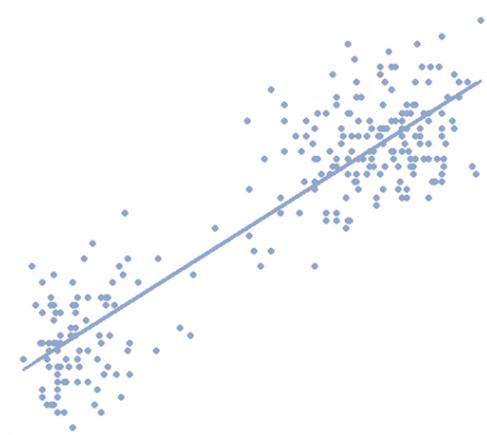
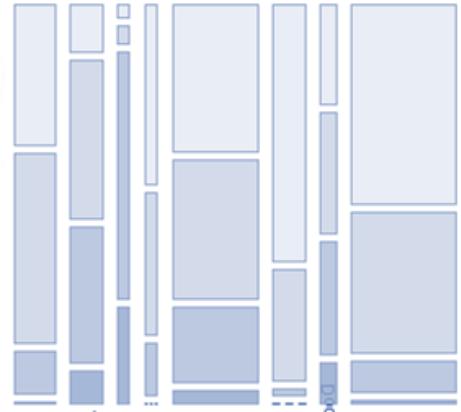
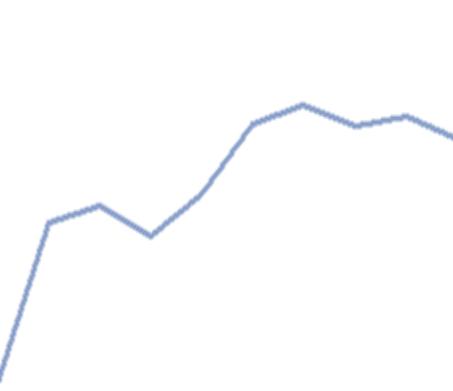
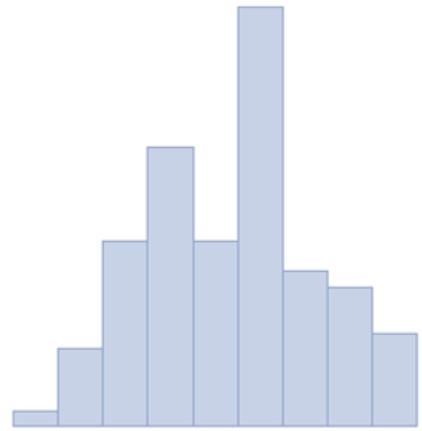
Data Visualization

Visual data representation



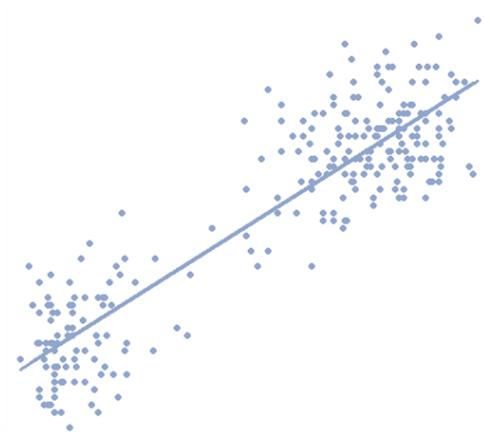
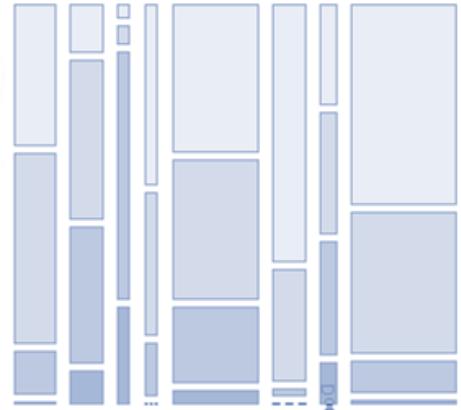
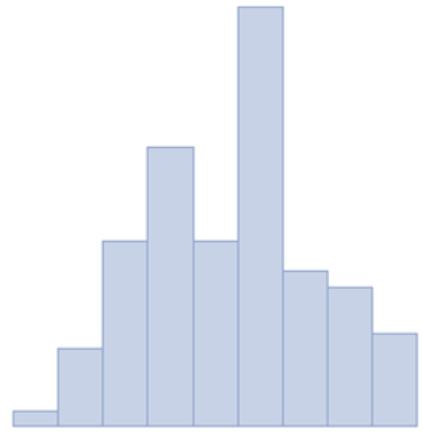
Data Visualization

Visual data representation
Human pattern recognition



Data Visualization

Visual data representation
Human pattern recognition
Map dimensions to visual

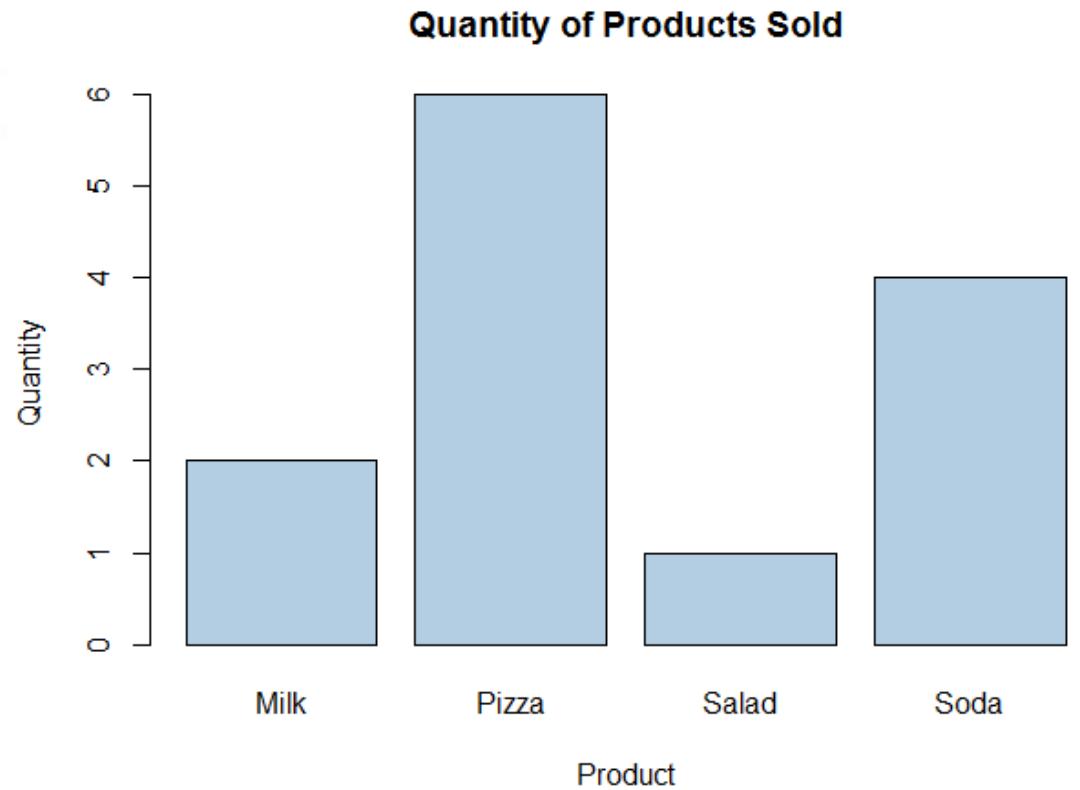


Data Visualization

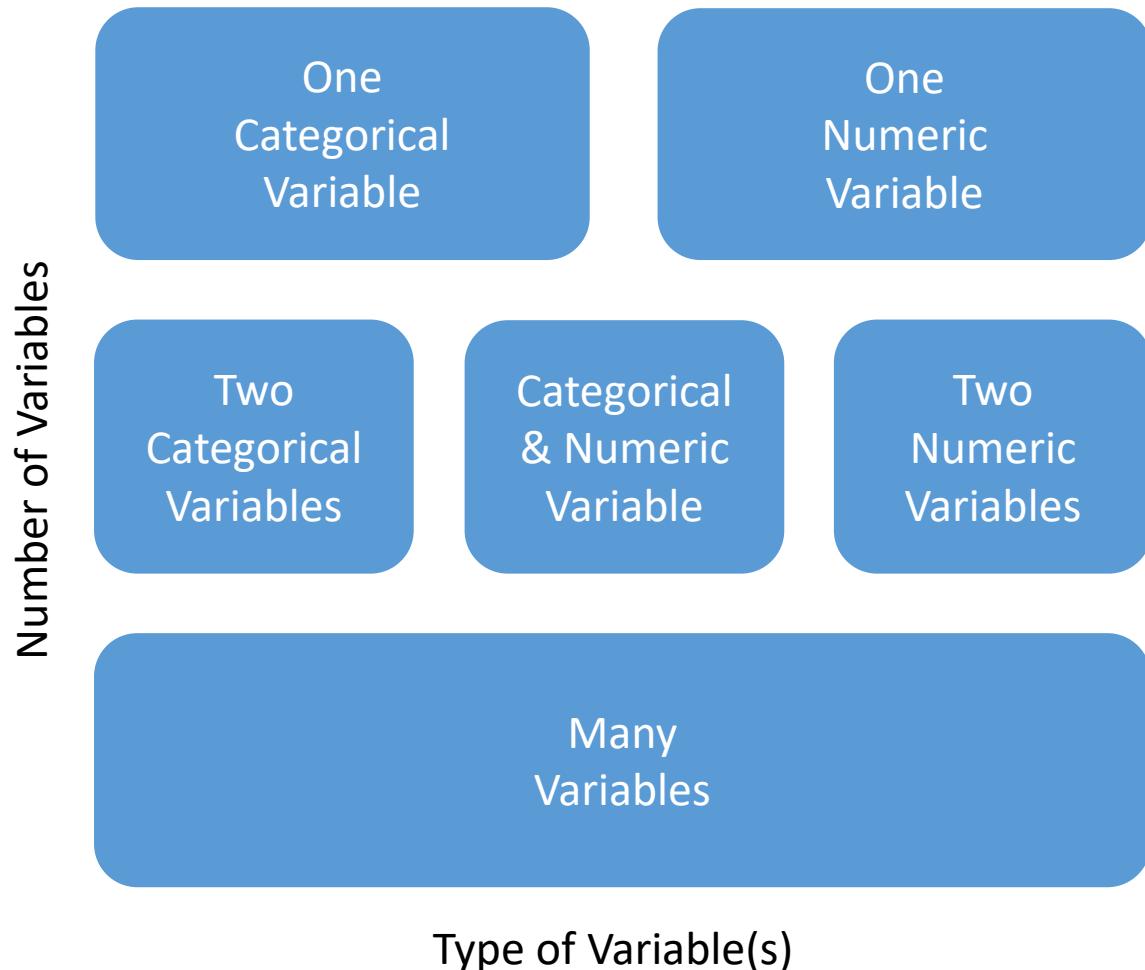
ID	Date	Customer	Product	Quantity
1	2015-08-27	John	Pizza	2
2	2015-08-27	John	Soda	2
3	2015-08-27	Jill	Salad	1
4	2015-08-27	Jill	Milk	1
5	2015-08-28	Miko	Pizza	3
6	2015-08-28	Miko	Soda	2
7	2015-08-28	Sam	Pizza	1
8	2015-08-28	Sam	Milk	1

Data Visualization

ID	Date	Customer	Product	Quantity
1	2015-08-27	John	Pizza	2
2	2015-08-27	John	Soda	2
3	2015-08-27	Jill	Salad	1
4	2015-08-27	Jill	Milk	1
5	2015-08-28	Miko	Pizza	3
6	2015-08-28	Miko	Soda	2
7	2015-08-28	Sam	Pizza	1
8	2015-08-28	Sam	Milk	1



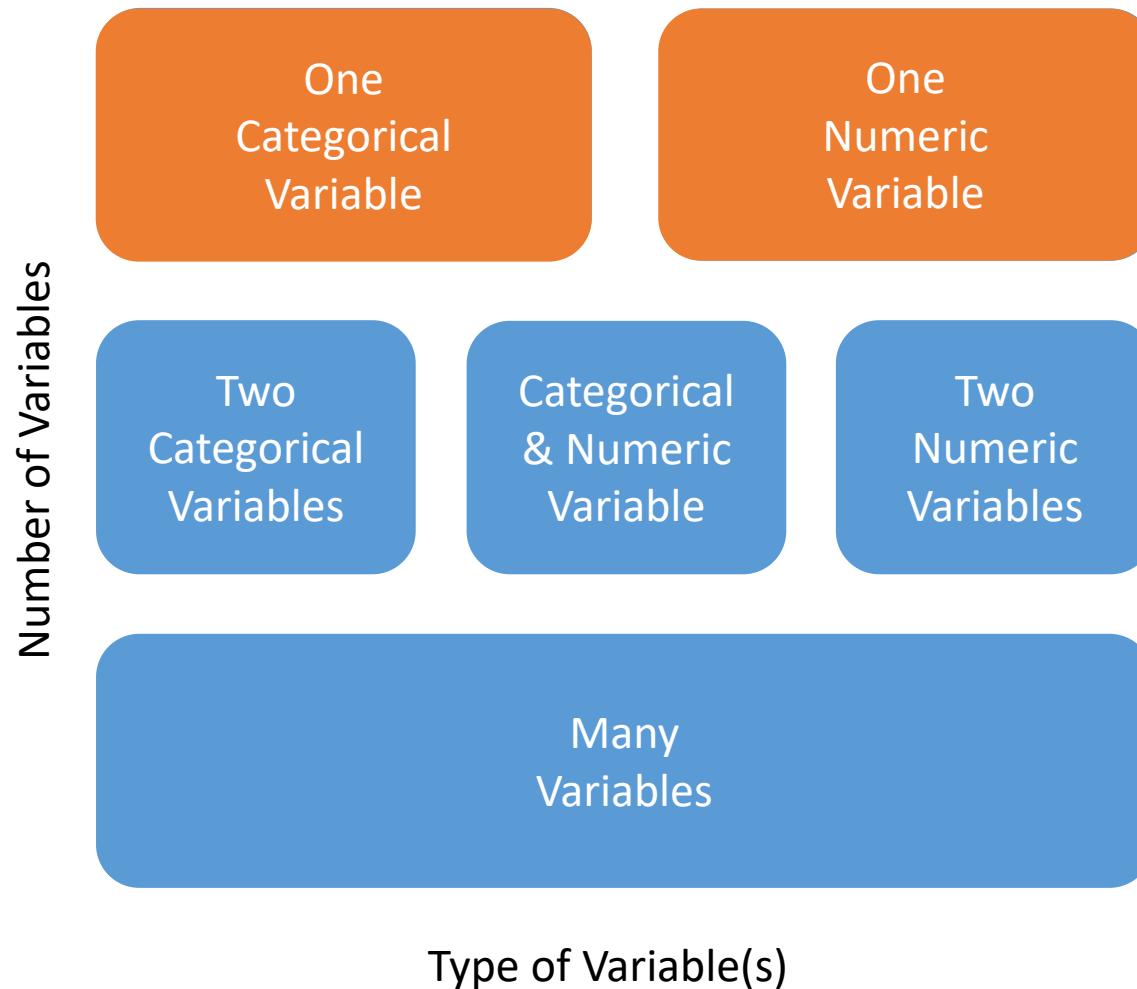
Types of Analysis



Code Demo

Visualizing One Variable

Visualizing One Variable



Analyzing One Categorical Variable

Frequency

Movies by Genre		
Genre	Frequency	Percentage
Action	612	9%
Adventure	496	7%
Animation	168	2%
Comedy	1281	18%
Drama	1570	22%
Horror	269	4%
...

Analyzing One Categorical Variable

Frequency
Proportion

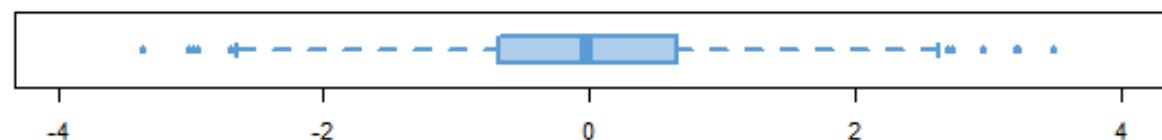
Movies by Genre		
Genre	Frequency	Percentage
Action	612	9%
Adventure	496	7%
Animation	168	2%
Comedy	1281	18%
Drama	1570	22%
Horror	269	4%
...

Visualizing One Numeric Variable

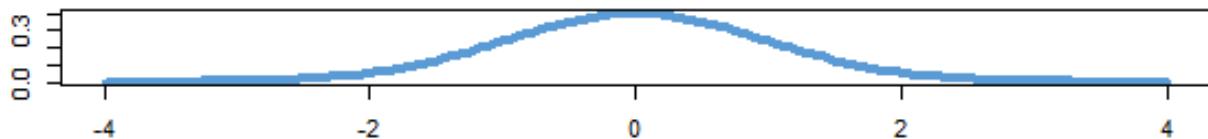
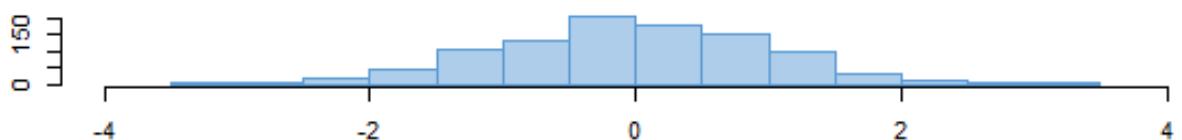
Location



Spread



Shape





PROD. NO.
SCENE

TAKE

ROLL

Open Movies Database

Movies								
Title	Year	Rating	Runtime (minutes)	Genre	Critic Score	Box Office (\$M)	Awards	Distribution
The Whole Nine Yards	2000	R	98	Comedy	45%	57.3	No	Domestic
Cirque du Soleil	2000	G	39	Family	45%	13.4	Yes	Domestic
Gladiator	2000	R	155	Action	76%	187.3	Yes	International
Dinosaur	2000	PG	82	Family	65%	135.6	Yes	Domestic
Big Momma's House	2000	PG-13	99	Comedy	30%	0.5	Yes	International



PROD. NO.
SCENE

TAKE

ROLL





1. How does the count of movies compare by rating category?
2. What proportion of movies have won awards?
3. What does the distribution of movie runtimes look like?

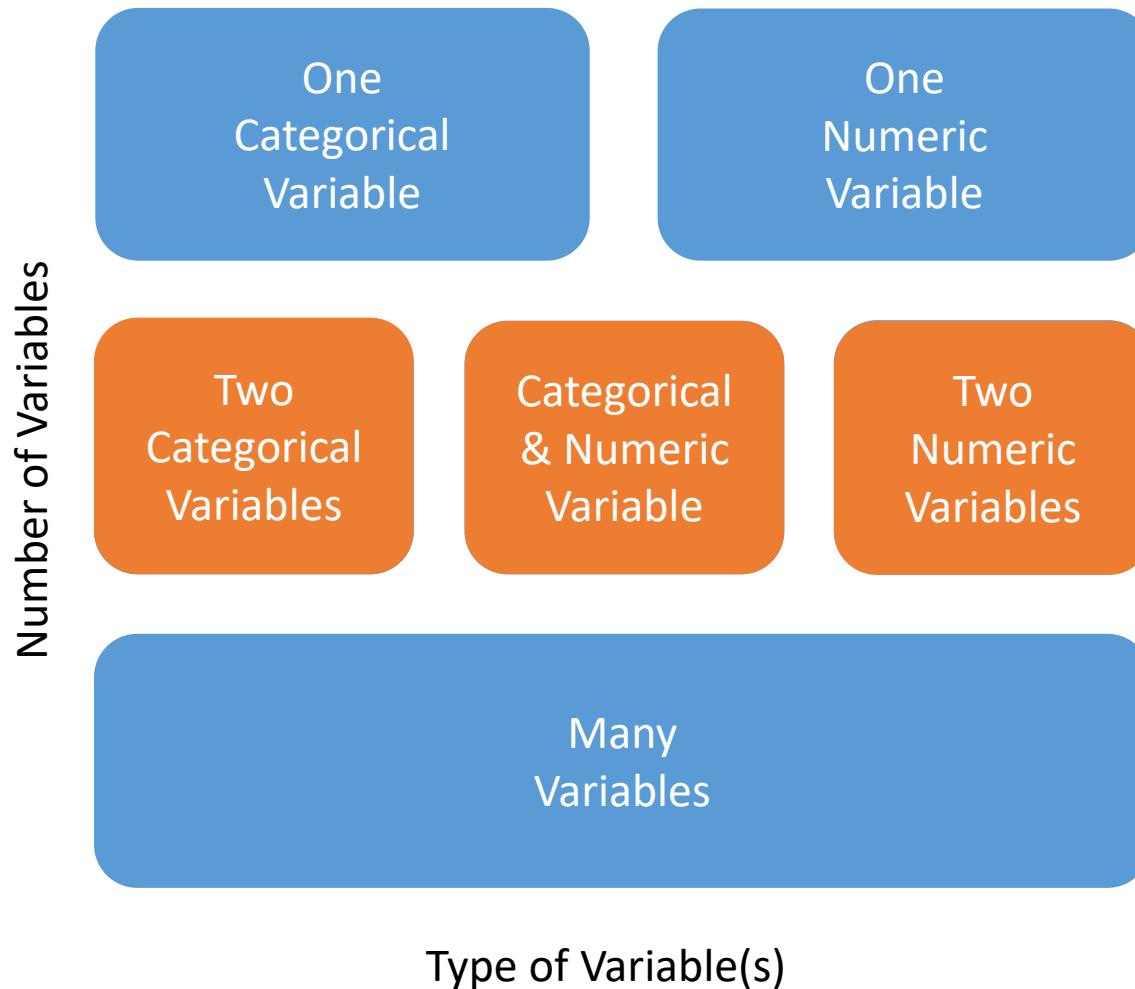
Code Demo

(Base graphics system)



Visualizing Two Variables

Visualizing Two Variables



Analyzing Two Categorical Variables

Joint frequency

Movies by Genre and Rating					
Genre	G	PG	PG-13	R	Total
Action	2	70	311	229	612
Adventure	44	179	209	64	496
Animation	43	111	8	6	168
Comedy	45	258	472	506	1218
Drama	12	136	586	836	1570
Family	38	181	10	1	230
...
Total	230	1207	2686	3058	7181

Analyzing Two Categorical Variables

Joint frequency
Contingency table

Movies by Genre and Rating					
Genre	G	PG	PG-13	R	Total
Action	2	70	311	229	612
Adventure	44	179	209	64	496
Animation	43	111	8	6	168
Comedy	45	258	472	506	1218
Drama	12	136	586	836	1570
Family	38	181	10	1	230
...
Total	230	1207	2686	3058	7181

Analyzing Two Categorical Variables

Joint frequency

Contingency table

Marginal frequency

Movies by Genre and Rating					
Genre	G	PG	PG-13	R	Total
Action	2	70	311	229	612
Adventure	44	179	209	64	496
Animation	43	111	8	6	168
Comedy	45	258	472	506	1218
Drama	12	136	586	836	1570
Family	38	181	10	1	230
...
Total	230	1207	2686	3058	7181

Analyzing Two Categorical Variables

Joint frequency

Contingency table

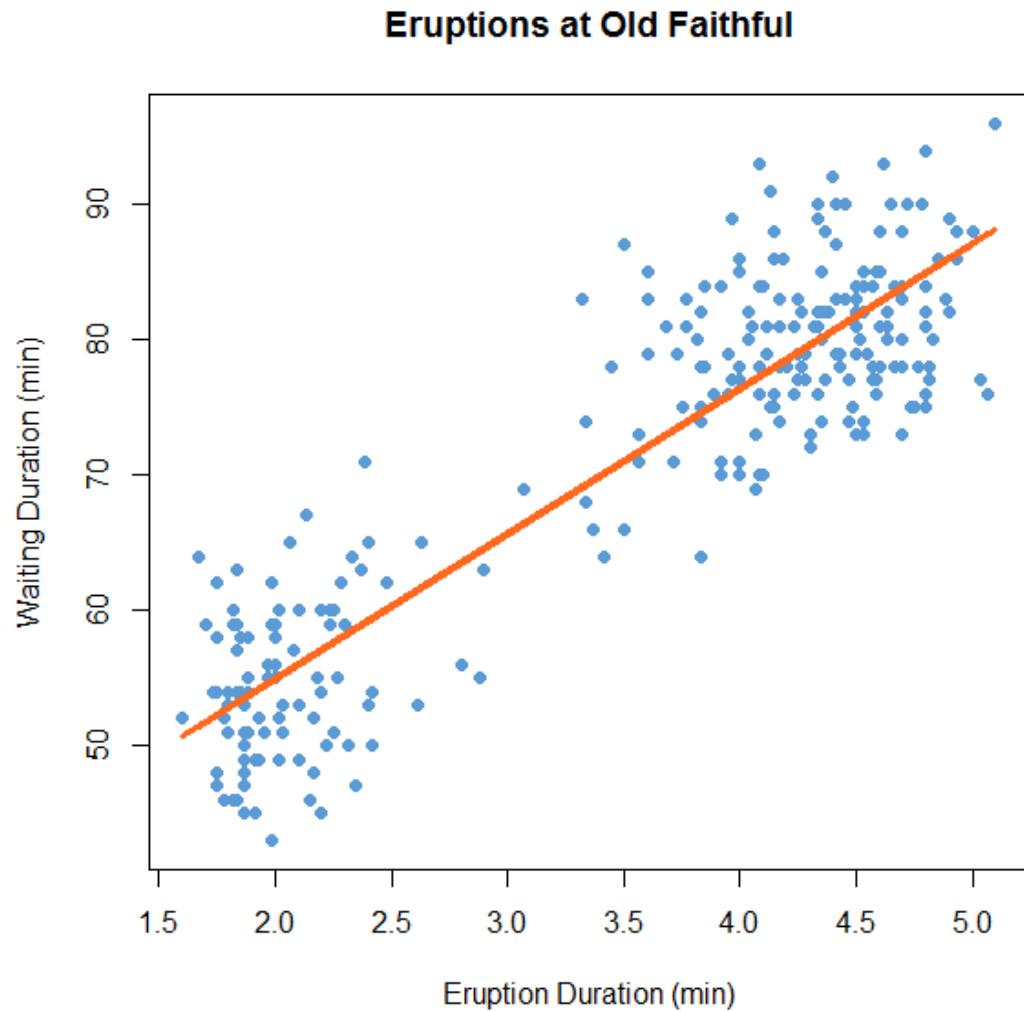
Marginal frequency

Relative frequency

Movies by Genre and Rating					
Genre	G	PG	PG-13	R	Total
Action	0.001	0.010	0.043	0.032	0.086
Adventure	0.006	0.025	0.029	0.009	0.069
Animation	0.006	0.015	0.001	0.001	0.023
Comedy	0.006	0.036	0.066	0.070	0.170
Drama	0.002	0.019	0.082	0.116	0.219
Family	0.005	0.025	0.001	0.001	0.033
...
Total	0.032	0.168	0.374	0.426	1.000

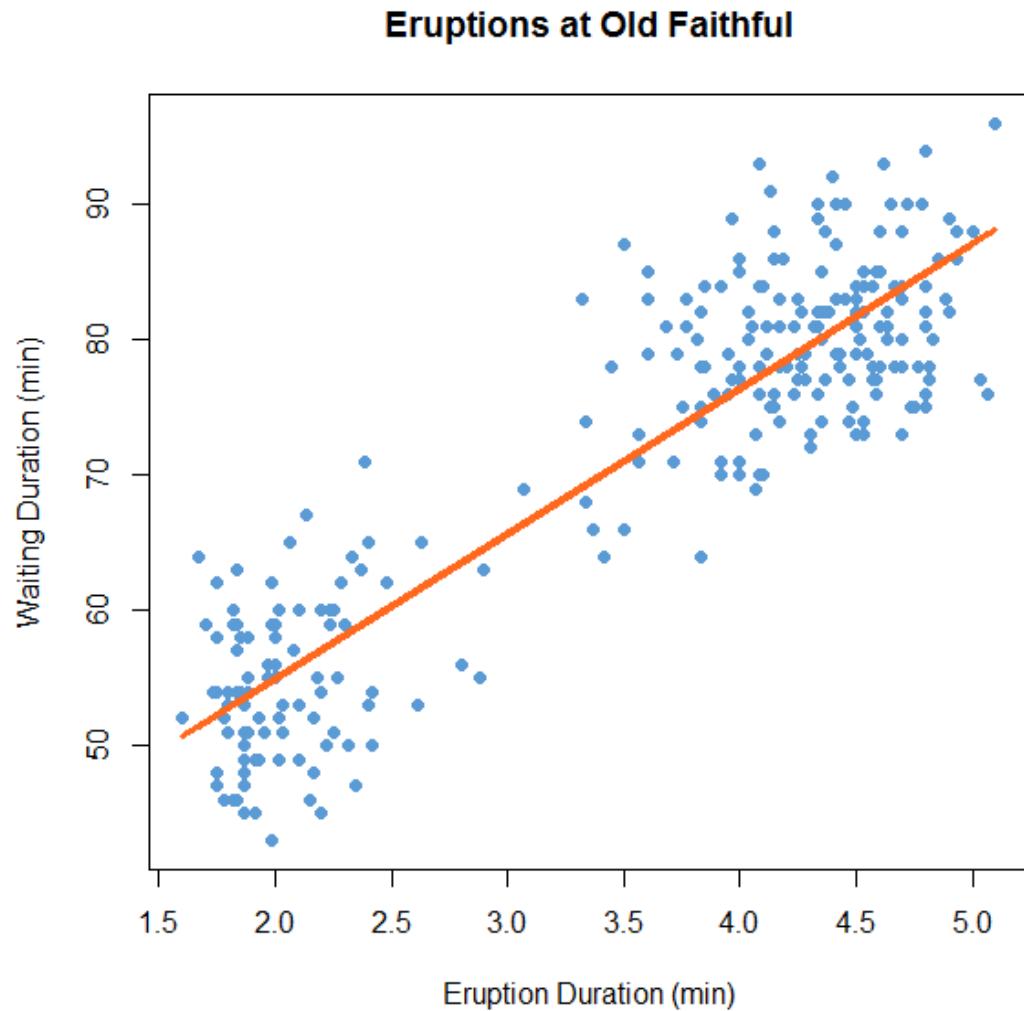
Visualizing Two Numeric Variables

Relationship



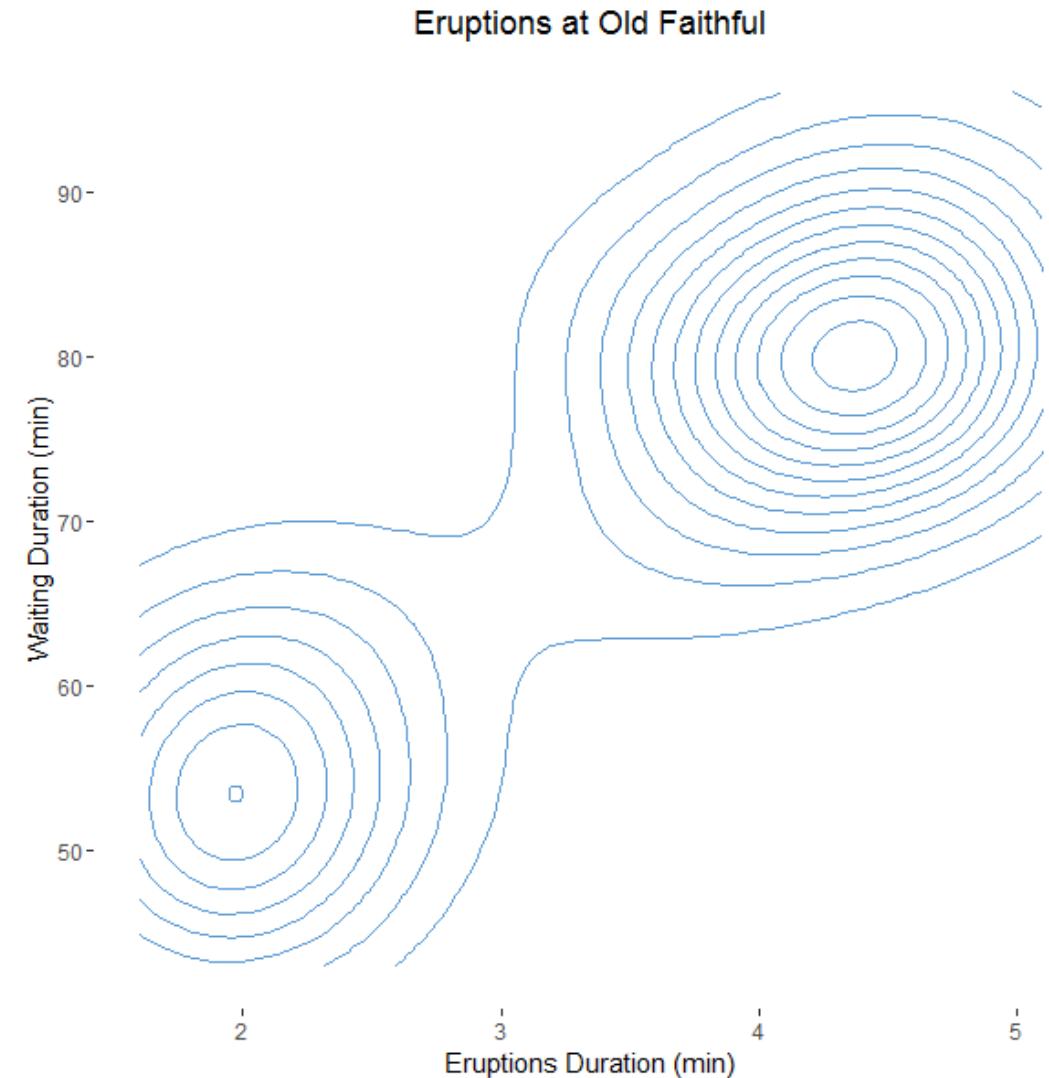
Visualizing Two Numeric Variables

Relationship
Correlation



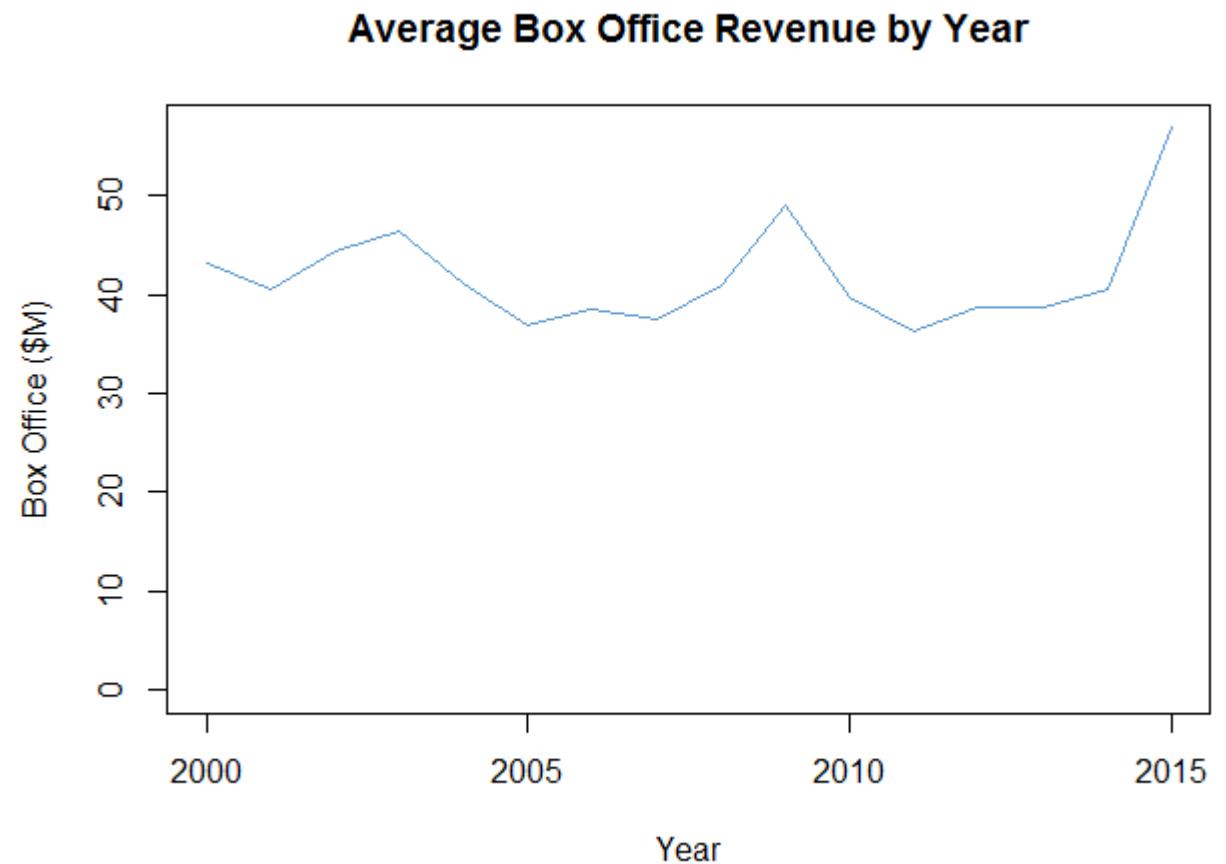
Visualizing Two Numeric Variables

Relationship
Correlation
Distribution



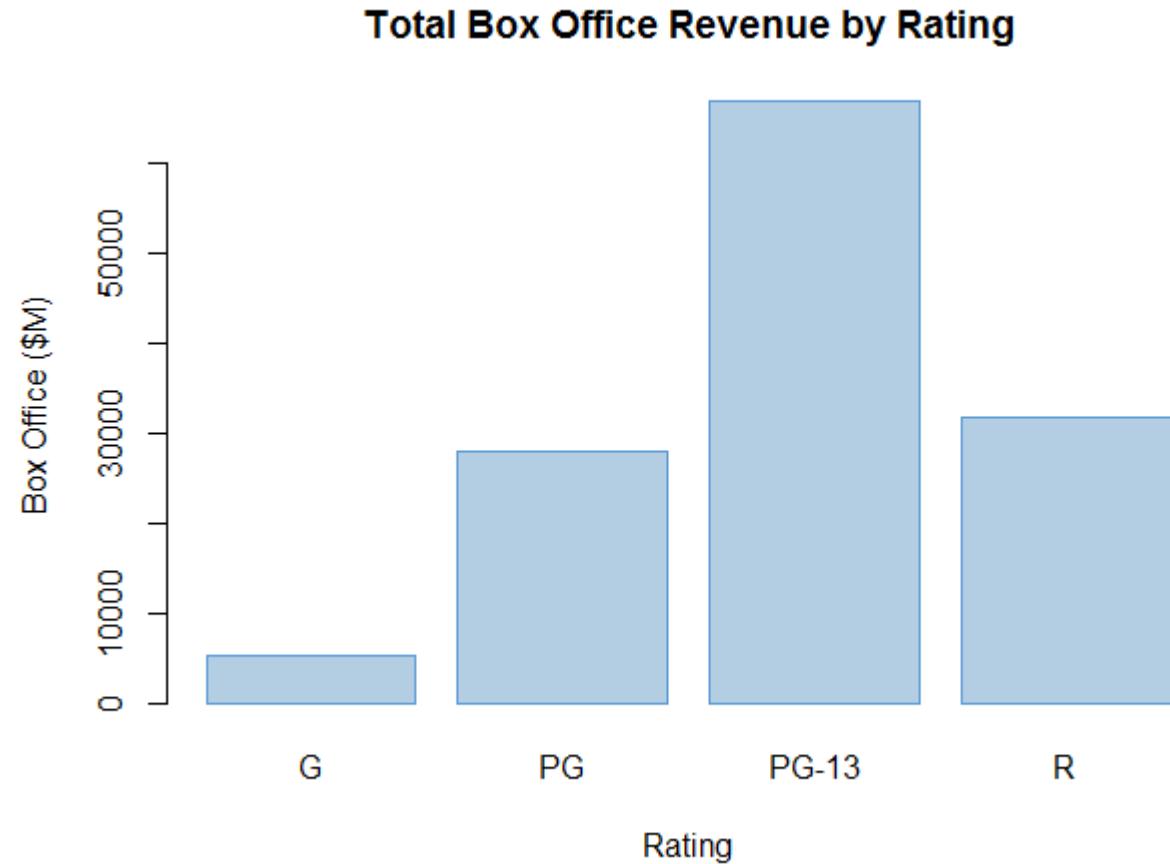
Visualizing Time Series Data

Values over time
Rate of change



Visualizing a Numeric Variable Grouped by a Categorical Variable

Aggregate
Grouped
Comparison





1. How many award-winning movies by rating category?
2. How are runtime and box-office revenue related?
3. Has the average box-office revenue changed over time?
4. How does box office compare by rating category?

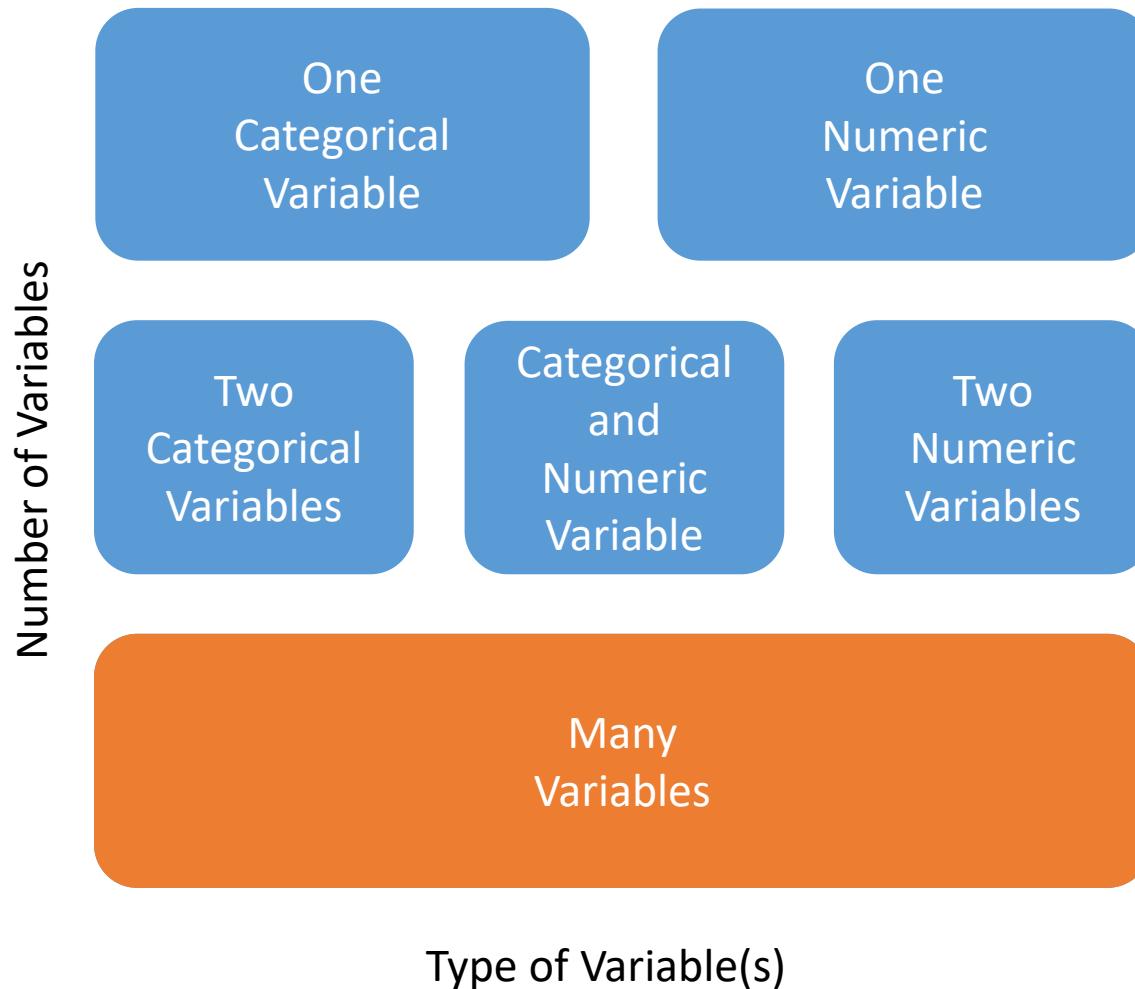
Code Demo

(Lattice)

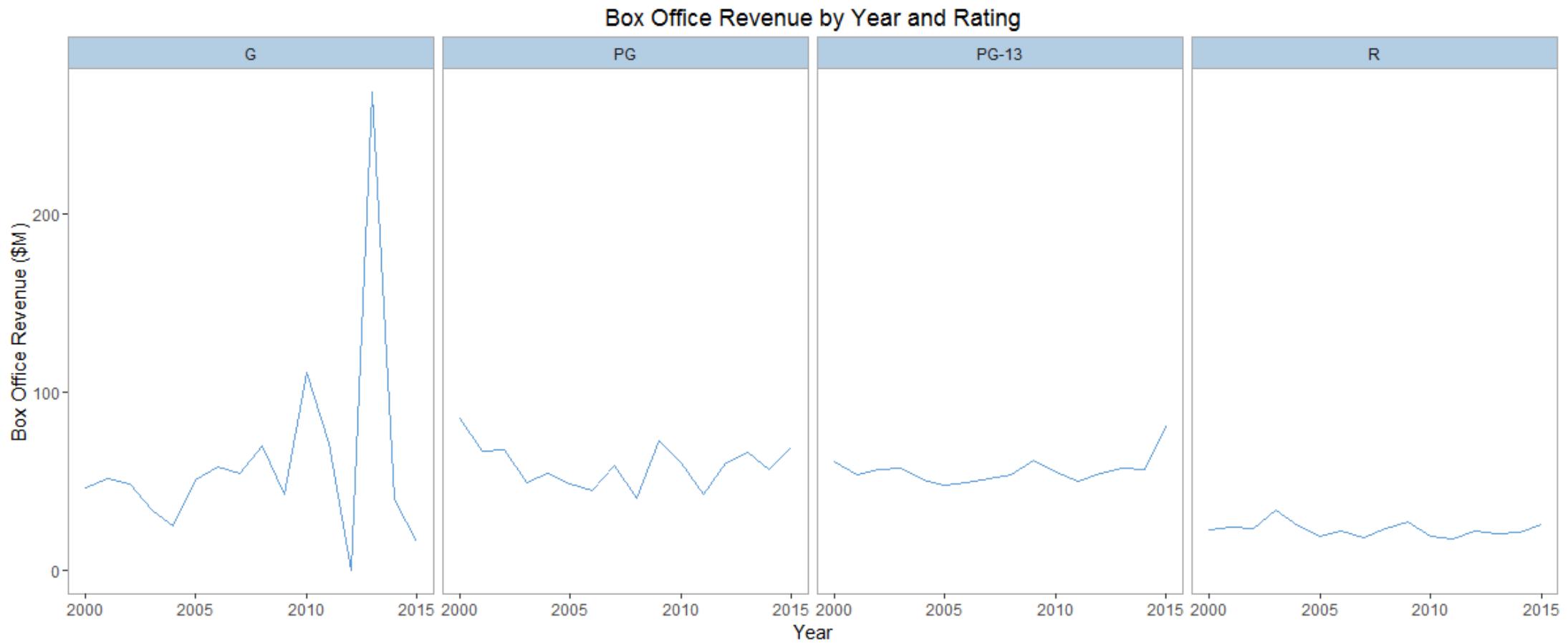


Visualizing Many Variables

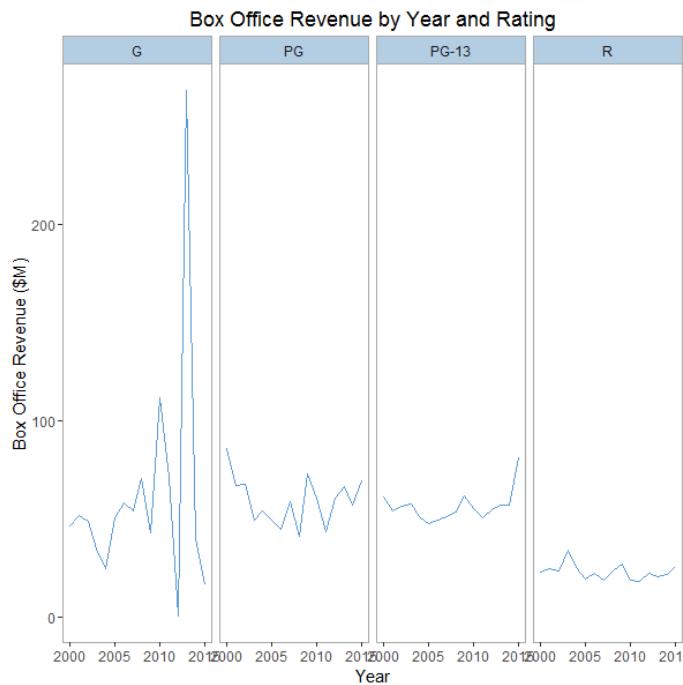
Visualizing Many Variables



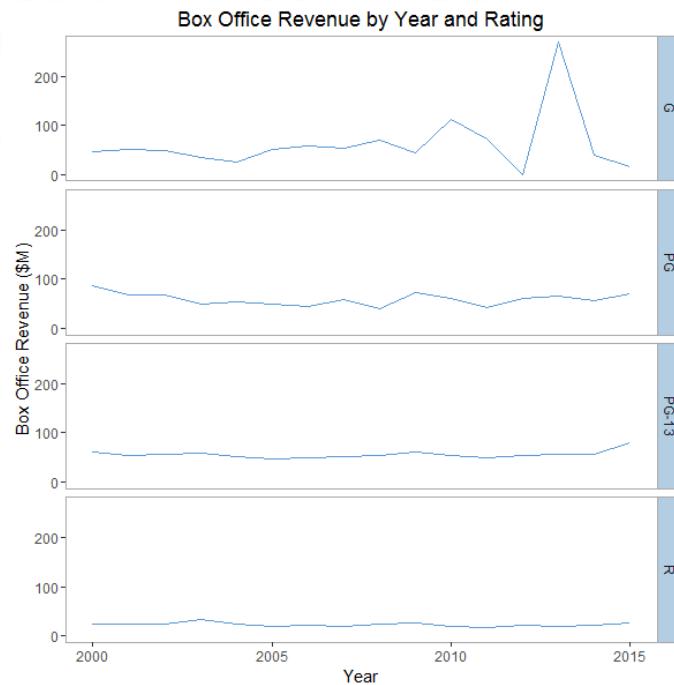
Faceting



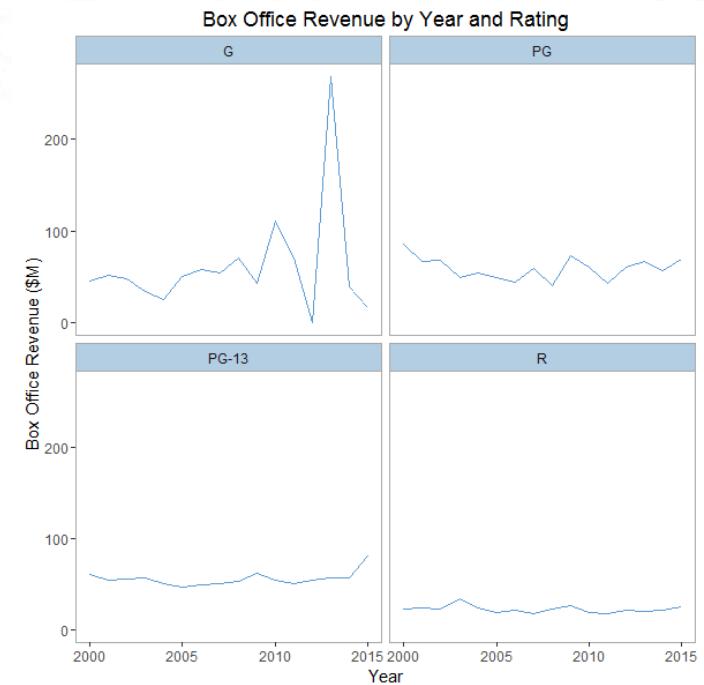
Faceting Orientation



Horizontal

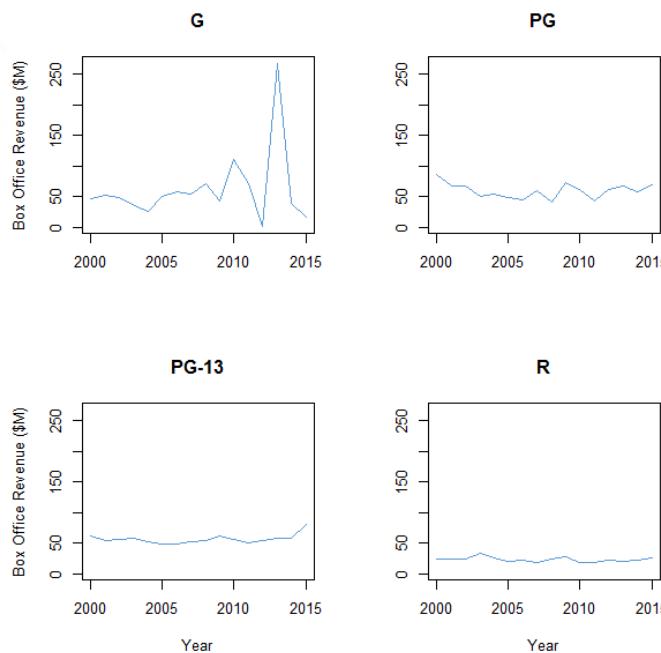


Vertical

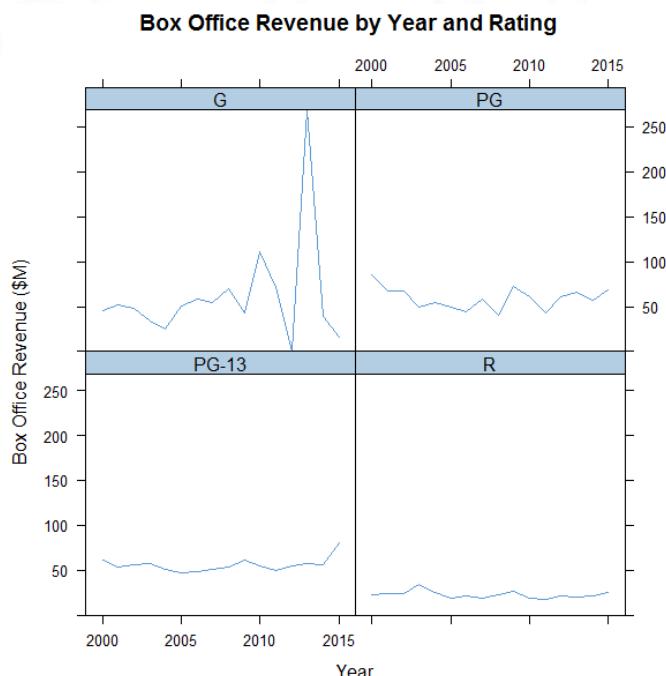


Wrapped

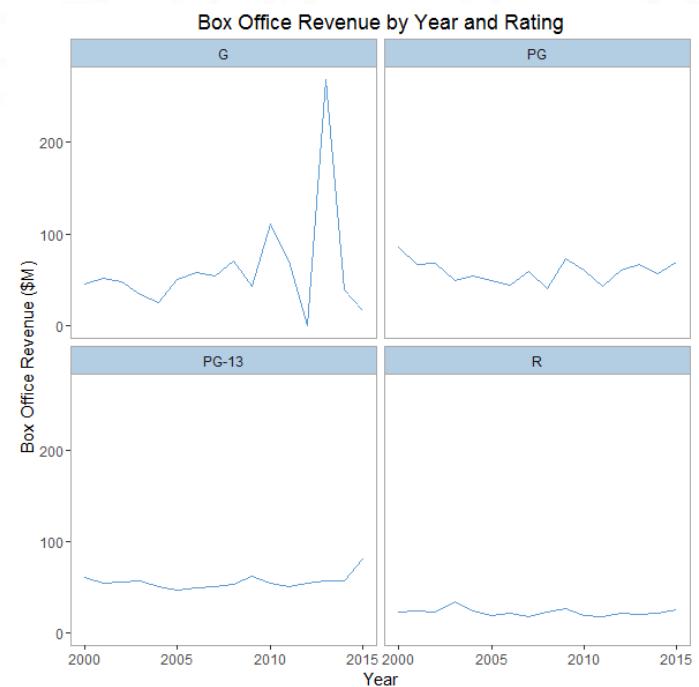
Faceting



Base



Lattice



ggplot2



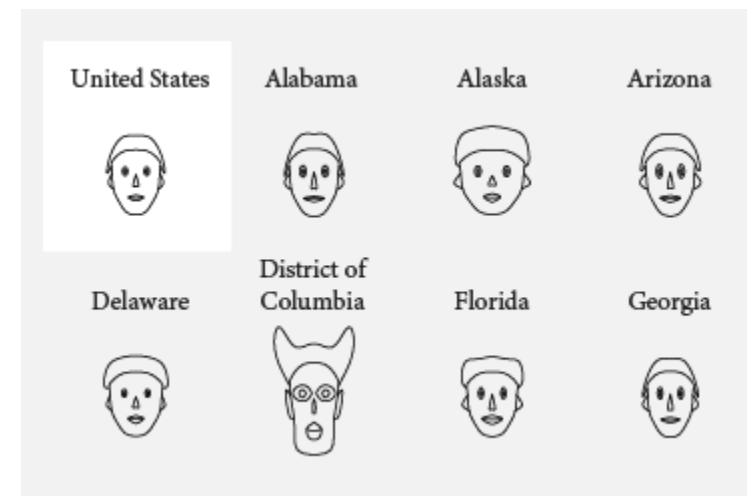
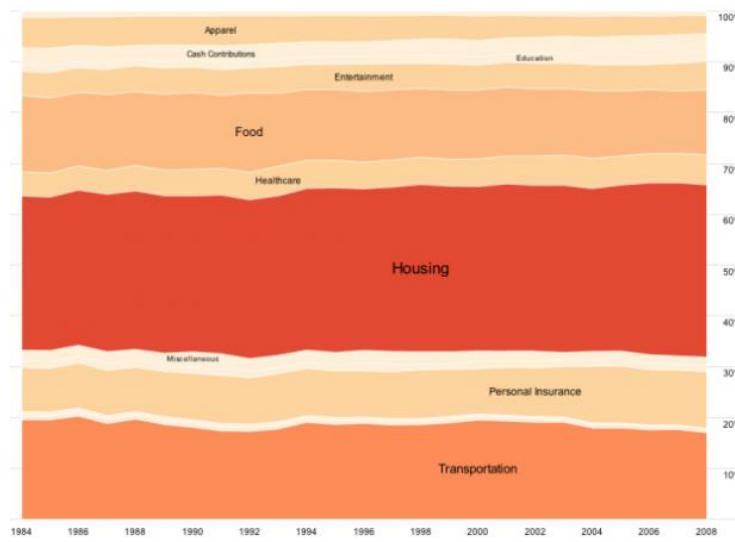
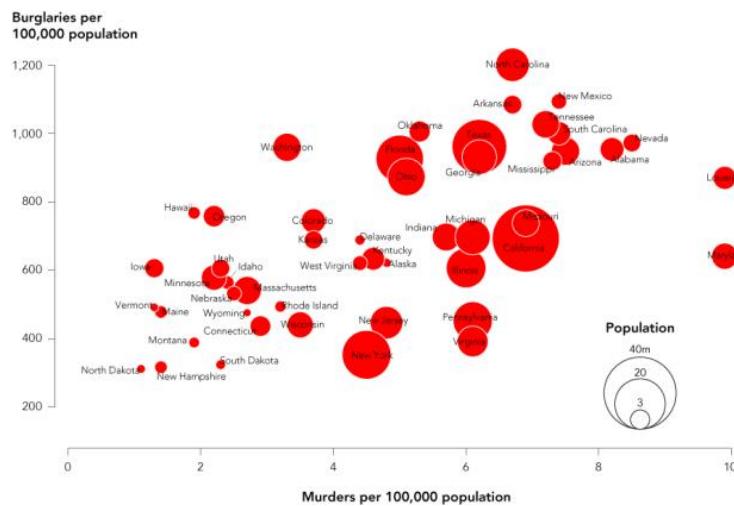
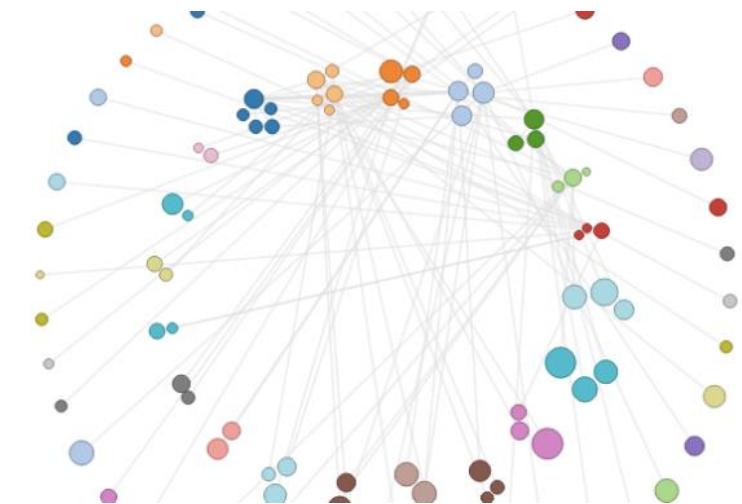
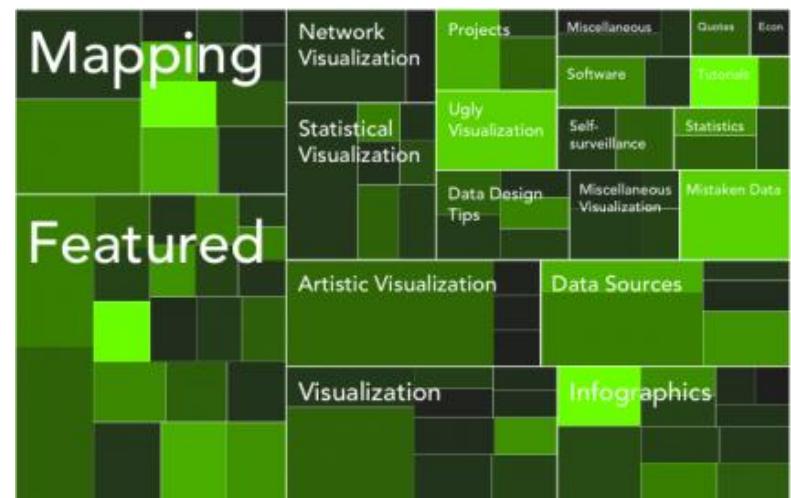
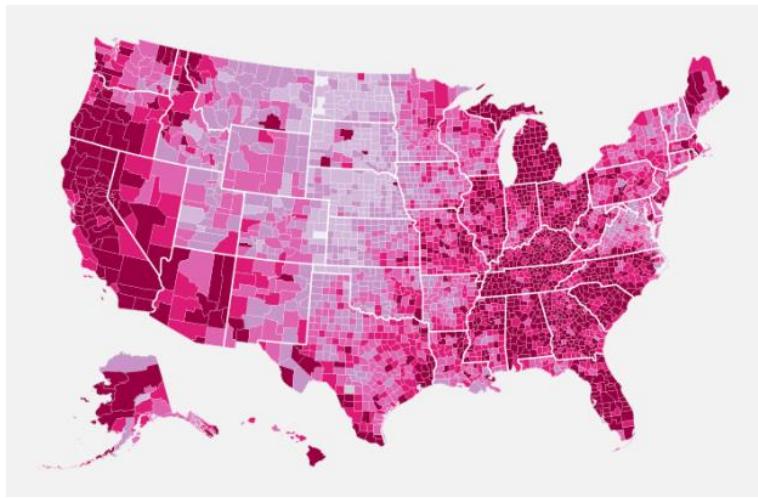
1. How many movies based on rating, award, and dist.?
2. How does box office compare across awards and rating?
3. How does critic score relate to box office by rating category?
4. How has box office revenue changed over time by rating?
5. How do runtime and critic score relate to box office?

Code Demo

(ggplot2)



Visualizing Other Types of Data



Source: Nathan Yau (www.flowingdata.com)



1. What is the box office revenue for films by country?
2. What is the box office revenue by continent and country?
3. Which actors have acted most frequently with one another?
4. Which movie plot words are associated with revenue?

Code Demo

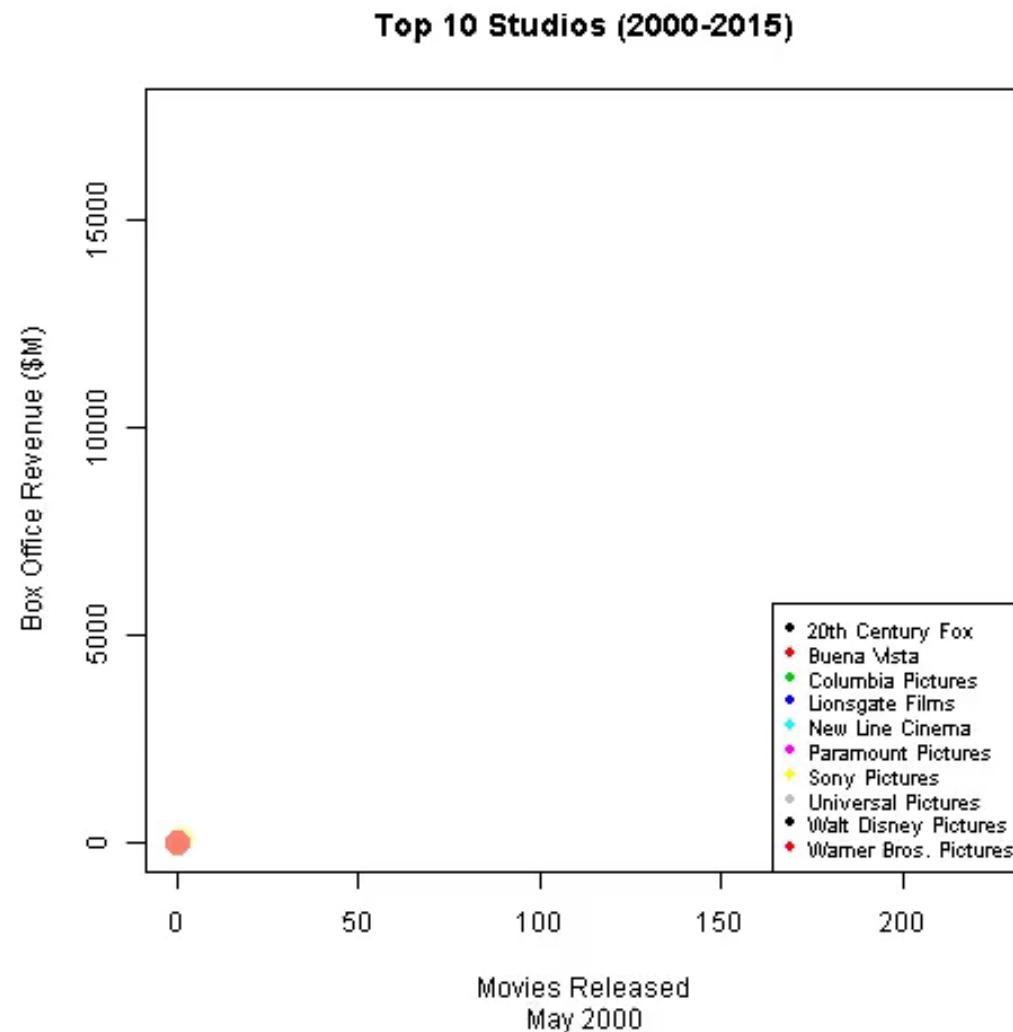


Beyond the Basics

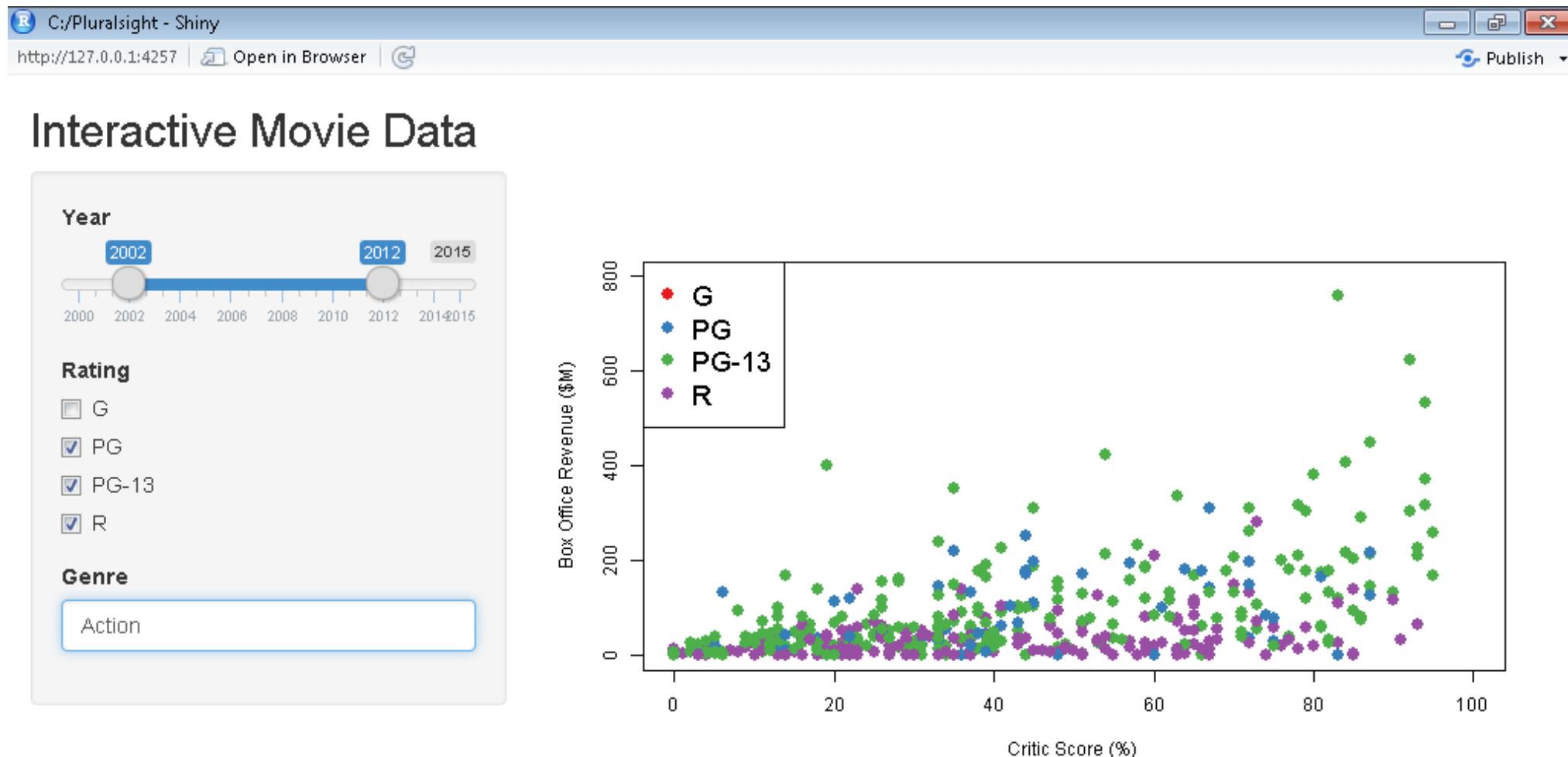


This is just the tip of the iceberg!

Animating Data Visualizations



Interactive Data Visualizations



How to Deploy to Production

Export charts (Rstudio)



Create documents (Markdown)



Create interactive reports (Shiny)

Deploy to Server (R Server)



Deploy to Cloud (Azure ML)





1. How have the top 10 studios grown over the past 15 years?
2. Allow me to explore the data to ask any question I might dream up on a random Tues.

... and I want it done today.

Code Demo



Where to Go Next...

R website: <http://www.cran.r-project.org>

RStudio: <https://www.rstudio.com>

Revolutions: <http://blog.revolutionanalytics.com>

Flowing Data: <http://flowingdata.com>

R-Blogger: <http://www.r-bloggers.com>

R-Seek: <http://rseek.org>



PLURALSIGHT

Data Visualization with R (3-part)

Data Science: The Big Picture

Data Science with R

Exploratory Data Analysis with R

Mastering Data Visualization with R



Matthew Renze
SOFTWARE CONSULTANT
@matthewrenze www.matthewrenze.com



www.pluralsight.com/authors/matthew-renze

News

2017-08-25 - Invitation to Speak at Devoxx Morocco

Very excited to announce that I've been invited to give a keynote in Casablanca at [Devoxx Morocco](#) in November. My keynote presentation will be on [Artificial Intelligence](#).



2017-08-16 - Invitation to Speak at Microsoft Ignite

I've been invited to speak at [Microsoft Ignite](#) in Orlando, Florida in September. This will be my first time speaking at Ignite. Talks will include both Data Science and Machine Learning with R.



Matthew is a data science consultant, author for [Pluralsight](#), international public speaker, a [Microsoft MVP](#), [ASPInsider](#), and open-source software contributor.

2017-08-14 - Dev on Fire Interview

Feedback

Very important to me!

One thing you liked?

One thing I could improve?



Conclusion

Conclusion

Introduction to R

Intro to Data Visualization

Types of Data Visualizations

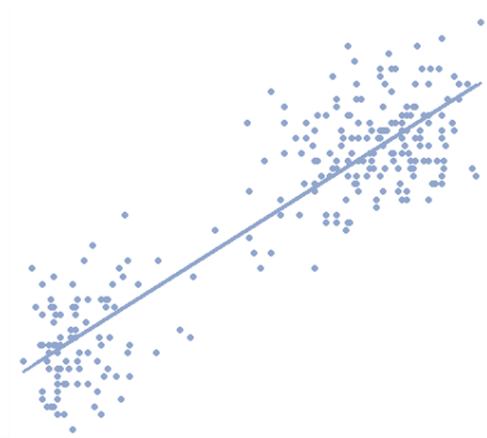
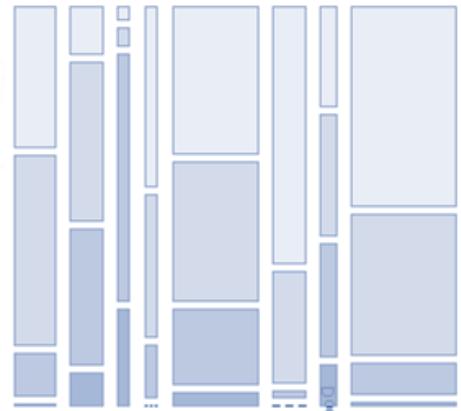
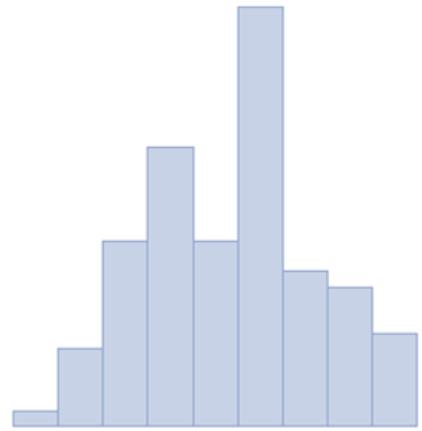
- Visualizing One Variable

- Visualizing Two Variables

- Visualizing Many Variables

- Visualizing Other Types of Data

Beyond the Basics



Thank You!

Matthew Renze

Data Science Consultant
Renze Consulting

Twitter: [@matthewrenze](https://twitter.com/matthewrenze)

Email: info@matthewrenze.com

Website: www.matthewrenze.com

