

# Quantitative Methoden in der Molekularbiologie

*thomas.rattei@univie.ac.at*

WS 2015/16

# Schedule

Oct 9	Numbers, significant digits and precision, diagrams, histogram
Oct 16	Protein quantification and mass spectrometry (Sampling, Normal distribution, Probabilities, Variability, Regression)
<i>Oct 23</i>	<i>Homework</i>
Oct 30	Planning an experiment (Sampling, design, simulation)
Nov 6+13	PCR, RT-QPCR (Normal distribution, probabilities, Null hypothesis, Testing, ANOVA)
<i>Nov 20</i>	<i>Homework</i>
Nov 27+ Dec 4	Human Microbiomes (Amplicons, Sequencing, Factorial designs, PCA, Covariance)
Dec 11+18	Microarrays/RNASeq: (clustering, classification, sensitivity, selectivity, ROC, multiple Testing)
Jan 8	Genetics (Genotype/Allele Frequencies, GWA, chi <sup>2</sup> -Test)
Jan 15	Recombinant DNA (sequence probabilistics, odds ratios, Bayes)
Jan 22	Exam (90 min.; written)
Jan 29	2D/3D Imaging (Image processing, pattern recognition, quantification)

# Literature

- McCleery, Watt, Hart: Introduction to Statistics for Biology
- Stephenson: Calculations for Molecular Biology and Biotechnology
- Krämer: So lügt man mit Statistik

# Contents

1. Significant numbers and scientific notation
2. Graphical representation of data

# Significant digits

- Numerals of a measurement representing actual limits of precision
- Surrounding zeros do not contribute to significant digits, regardless of notation
- E.g.: in 0.000 000 000 003 5 g, only 3 and 5 are significant
- Likewise in  $3.5000\text{e-}12$

# Problems:

- Significant numbers in:
  - a. 3 001 000 000 bp
  - b. 0.003 04 g
  - c. 0.000 210L

# Problems:

- Significant numbers in:
  - a. 3 001 000 000 bp 4
  - b. 0.003 04 g 3
  - c. 0.000 210L 2

# Rounding off significant digits in calculations

- The result can only be as precise as the least precise value
  - Adding or subtracting -> fewest significant digits to the right of the decimal
  - Multiplying or dividing -> fewest number of significant digits



# Problems:

- Significant numbers in:
  - a.  $0.2884 \text{ g} + 28.3 \text{ g}$
  - b.  $3.4 \text{ cm} * 8.115 \text{ cm}$
  - c.  $1.2 \text{ L} * 0.155 \text{ L}$

# Problems:

- Significant numbers in:
  - a.  $0.2884\text{ g} + 28.3\text{ g}$  (3 and) 1 decimal
  - b.  $3.4\text{ cm} * 8.115\text{ cm}$  2
  - c.  $1.2\text{ L} * 0.155\text{ L}$  2

# Exponents and scientific notation

- Decimal point right of leftmost nonzero
- Appropriate number of significant digits
- Exponent to basis 10

# Problems:

- Scientific notation of:
  - a. 3 001 000 000
  - b. 78
  - c.  $60.23 * 10^{22}$

# Problems:

- Scientific notation of:
  - a. 3 001 000 000      3.001e+9
  - b. 78      7.8e+1
  - c.  $60.23 * 10^{22}$       6.023e+23

# Problems:

- Decimal notation of:
  - a.  $4.37e+05$
  - b.  $2e+01$
  - c.  $23.4e+07$
  - d.  $3.2e-04$

# Problems:

- Decimal notation of:

a.	$4.37e+05$	437 000
b.	$2e+01$	20
c.	$23.4e+07$	234 000 000
d.	$3.2e-04$	0.000 32

# Problems:

- Calculating in scientific notation:
  - a.  $8e+04 + 7e+04$
  - b.  $2e+03 + 3e+01$
  - c.  $2.4e-03 - 3.7e-04$
  - d.  $3e+04 * 5e+02$
  - e.  $2e+03 * 6e-05$
  - f.  $8.2e-06 / 3.6e+04$



# Problems:

- Calculating in scientific notation:
  - a.  $8e+04 + 7e+04$        $1.5e+05$  or  $2e+05$
  - b.  $2e+03 + 3e+01$        $2e+03$
  - c.  $2.4e-03 - 3.7e-04$        $2.0e-03$
  - d.  $3e+04 * 5e+02$        $2e+07$
  - e.  $2e+03 * 6e-05$        $1e-01$
  - f.  $8.2e-06 / 3.6e+04$        $2.3e-10$

# Metric prefixes

- Use metric units
- Use prefixes: peta...atto ( $10^{15}$ .. $10^{-18}$ )

# Problems:

- Convert to metric units and prefixes:
  - a.  $6 \times 10^9$  bp in diploid human genome
  - b. 0.03 mg  $\rightarrow$  ng?
  - c. 0.0025 mL  $\rightarrow$   $\mu$ L?
  - d. 750 Da
  - e. 1.5 Å

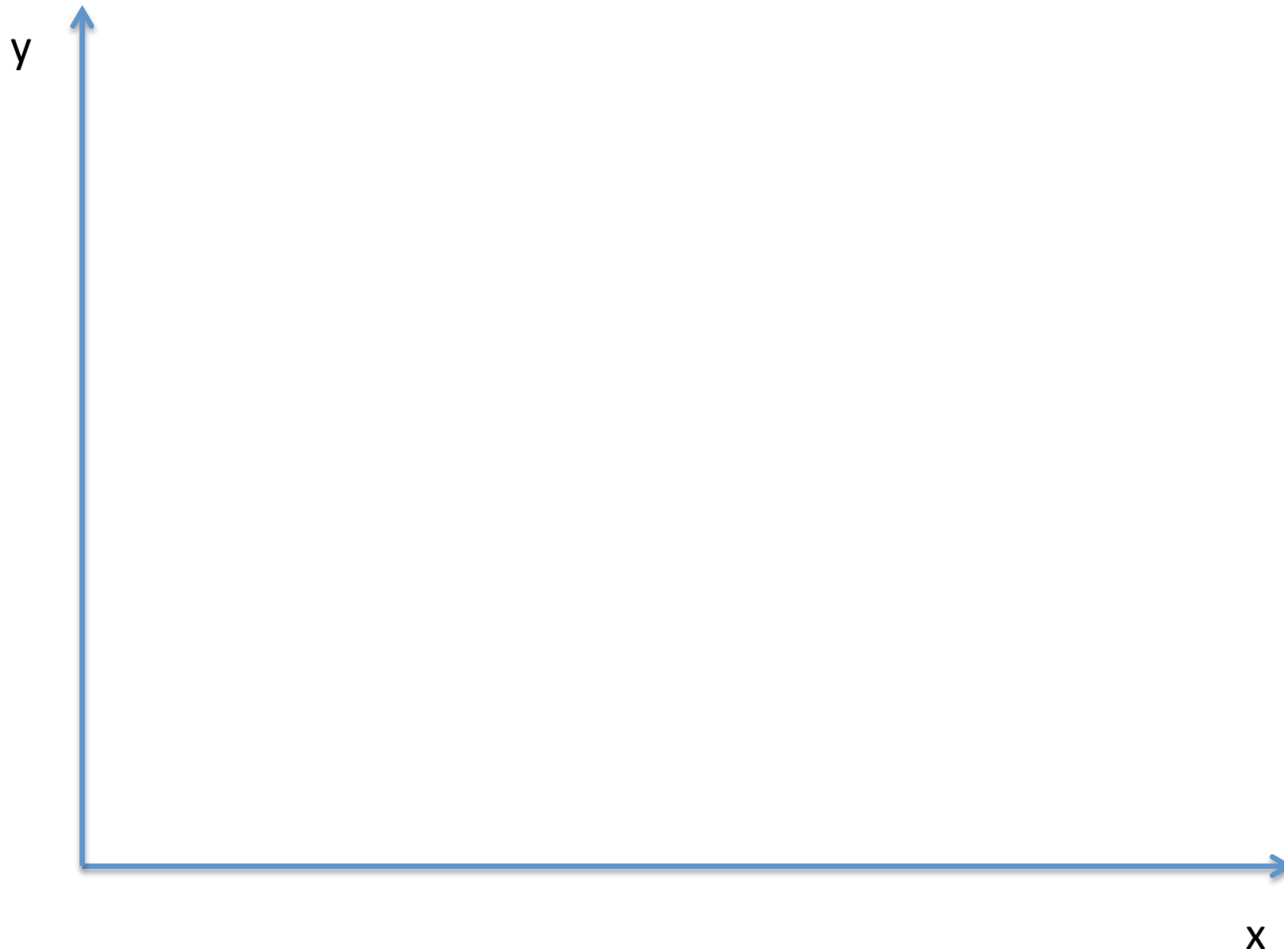
# Problems:

- Convert to metric units and prefixes:
  - a.  $6 \times 10^9$  bp in diploid human genome      6 Gbp
  - b. 0.03 mg  $\rightarrow$  ng?      30000 ng
  - c. 0.0025 mL  $\rightarrow$   $\mu$ L?      2.5  $\mu$ L
  - d. 750 Da      750 g/mol
  - e. 1.5 Å      0.15 nm

# Contents

1. Significant numbers and scientific notation
2. Graphical representation of data

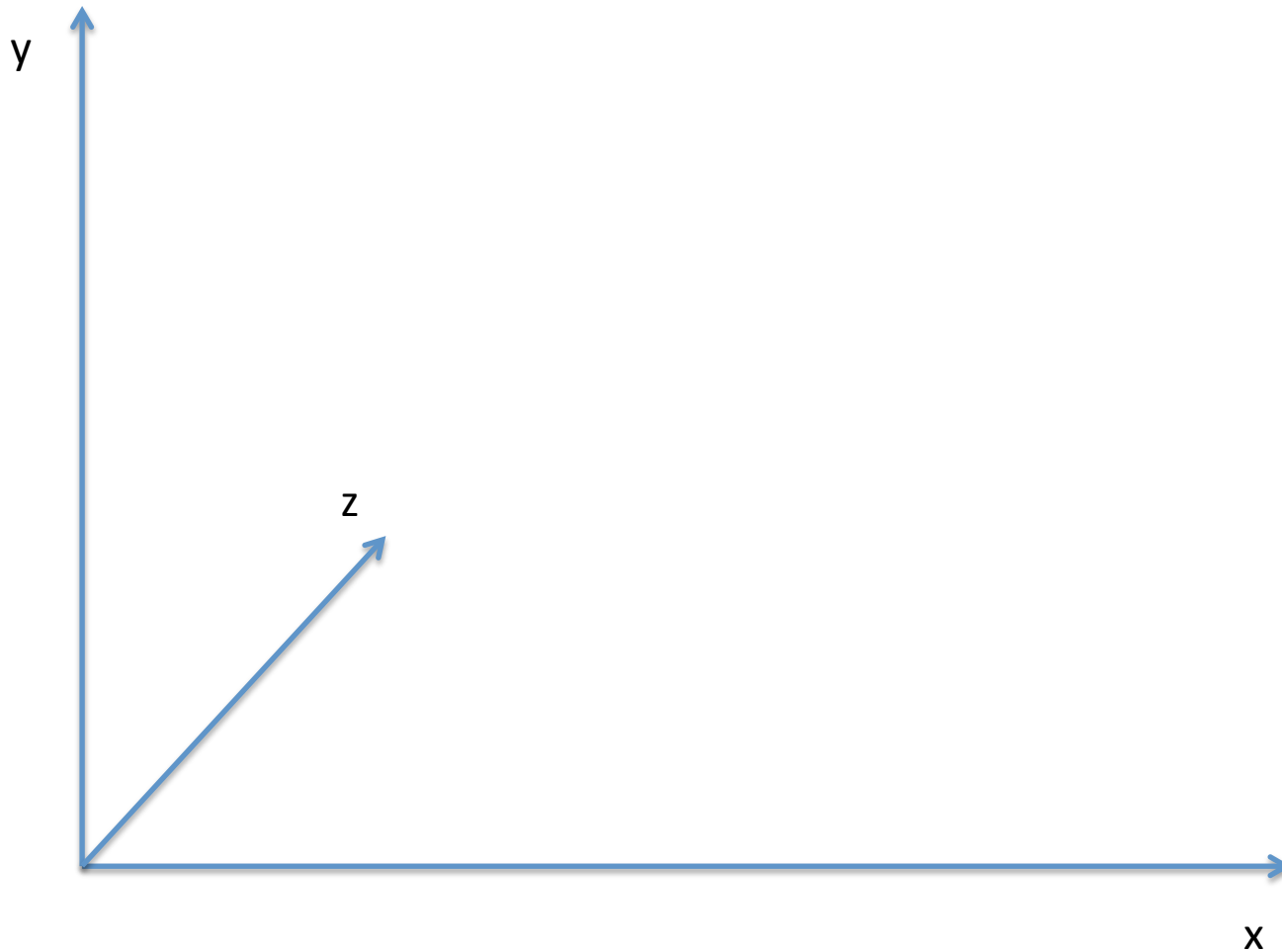
# 2D Plots



# Orientation and scaling of axes

- Axes
  - X: independent
  - Y: dependent
- Scaling:
  - No (x only as values)
  - Linear
  - Logarithmic

# 3D Plots





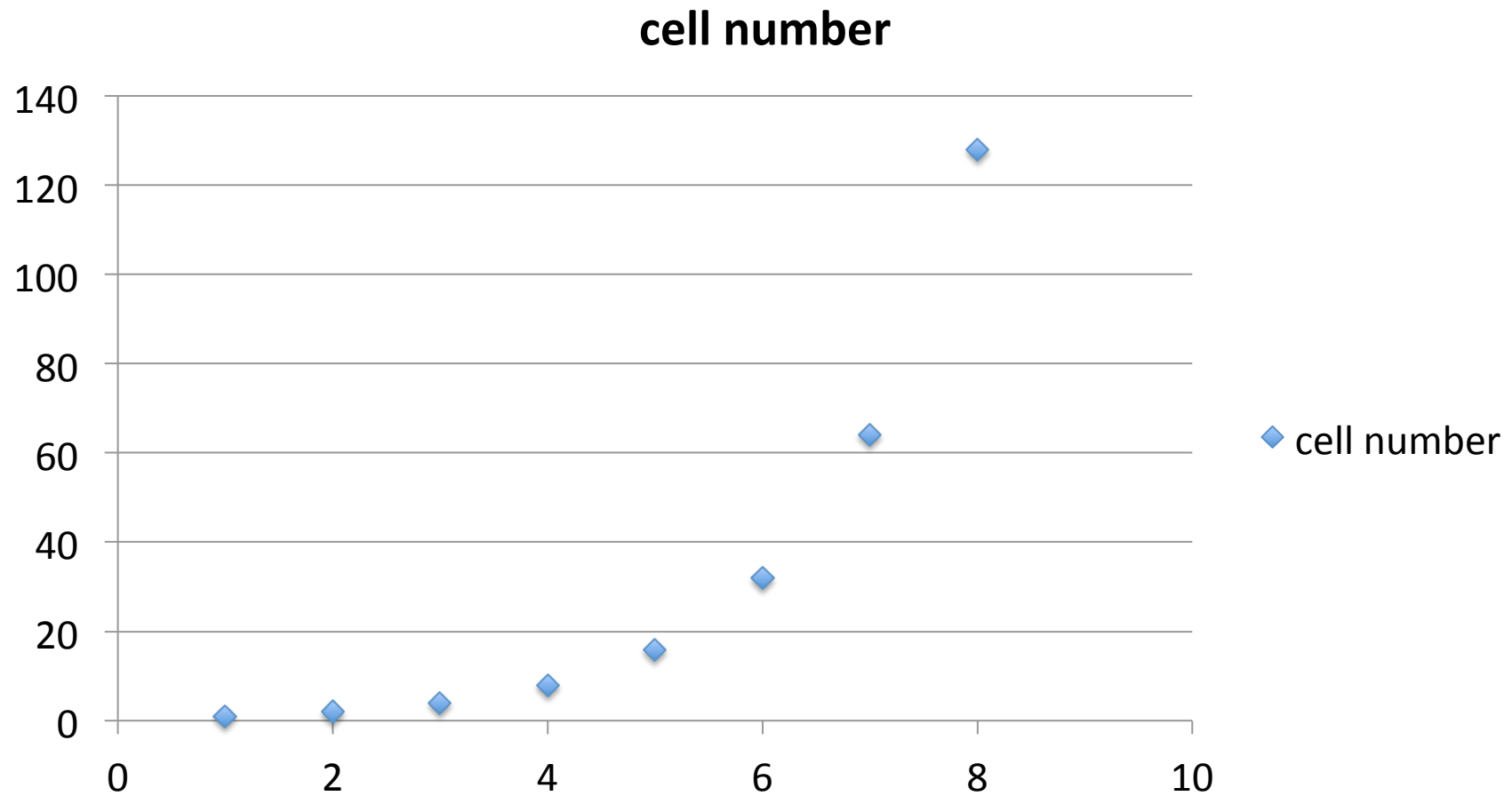
# Orientation and scaling of axes

- Axes
  - X: independent
  - Z: independent
  - Y: dependent
- Scaling:
  - No (x or/and z values only)
  - Linear
  - Logarithmic

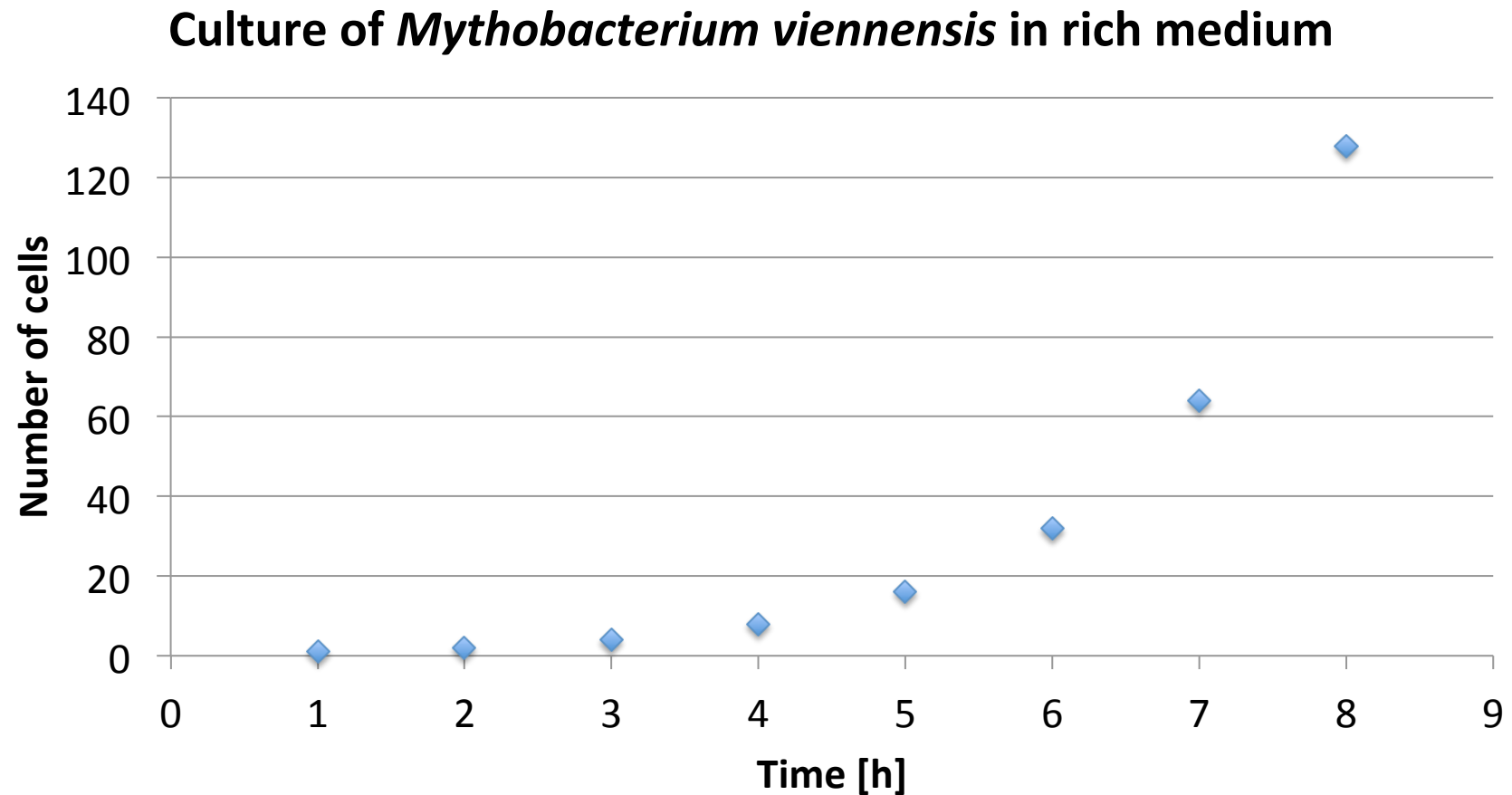
# Visualizing data in spreadsheets

Time [h]	cell number
1	1
2	2
3	4
4	8
5	16
6	32
7	64
8	128

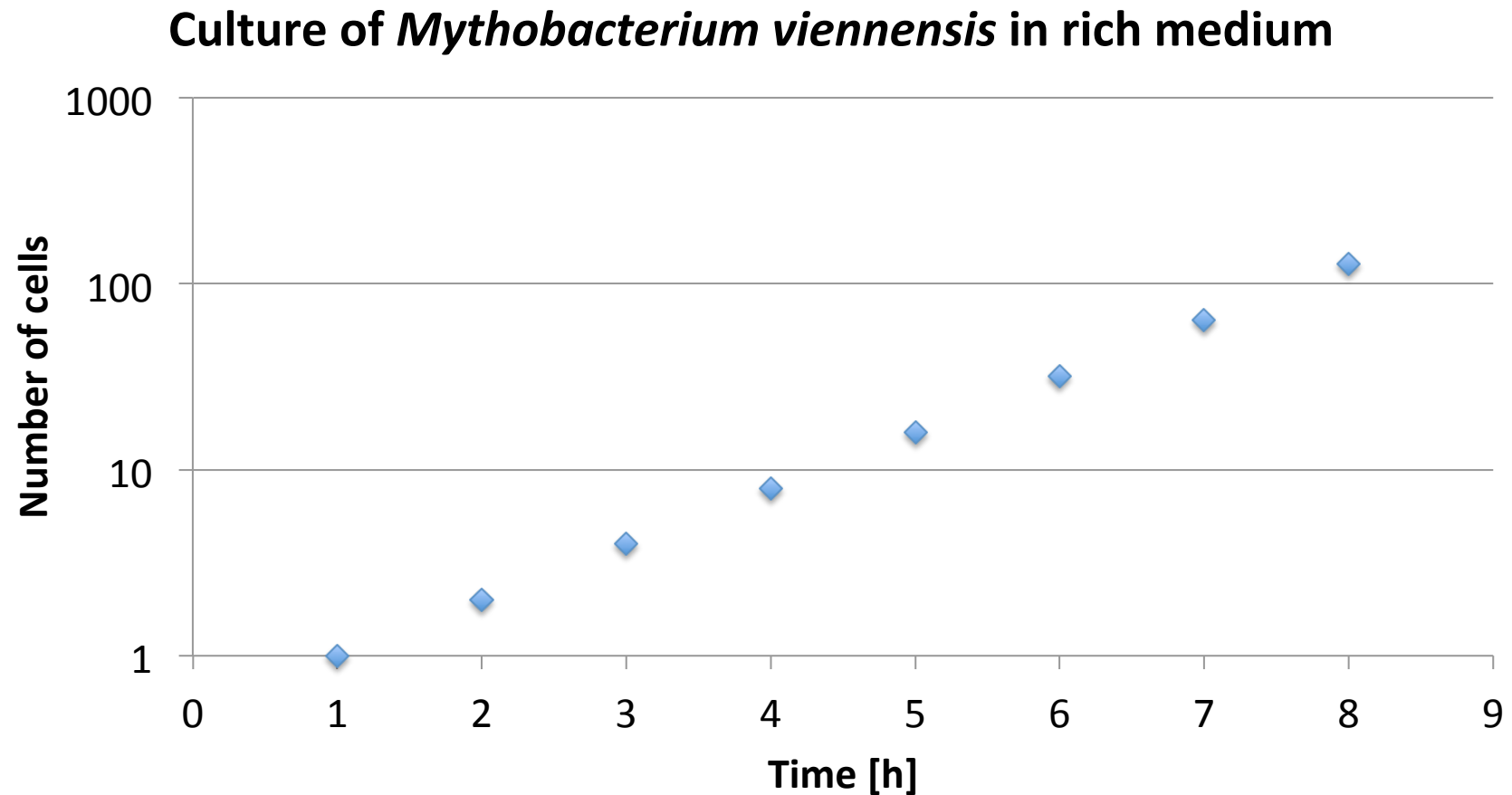
# Visualizing data in spreadsheets



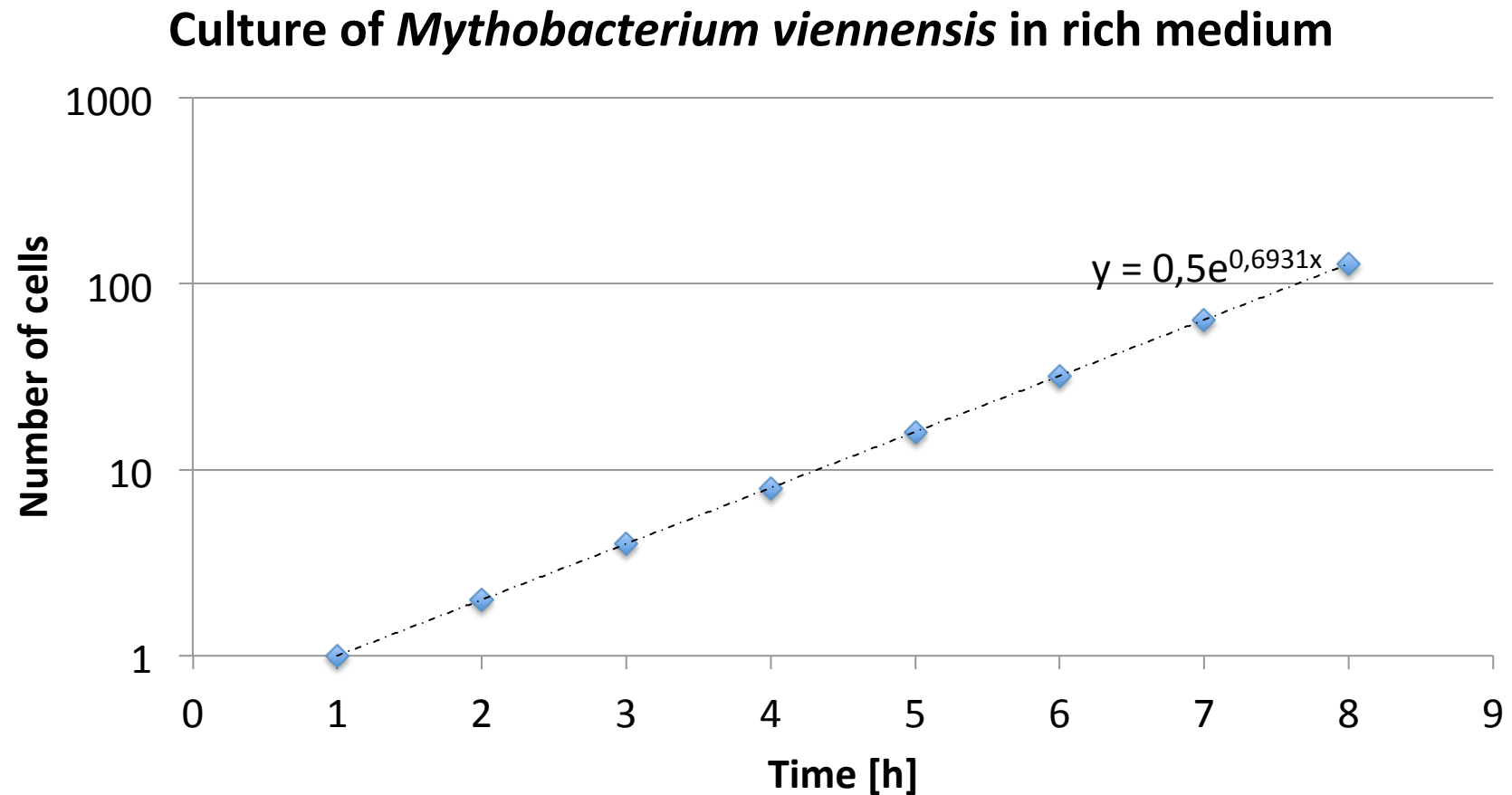
# Visualizing data in spreadsheets



# Visualizing data in spreadsheets



# Visualizing data in spreadsheets



# Histogram

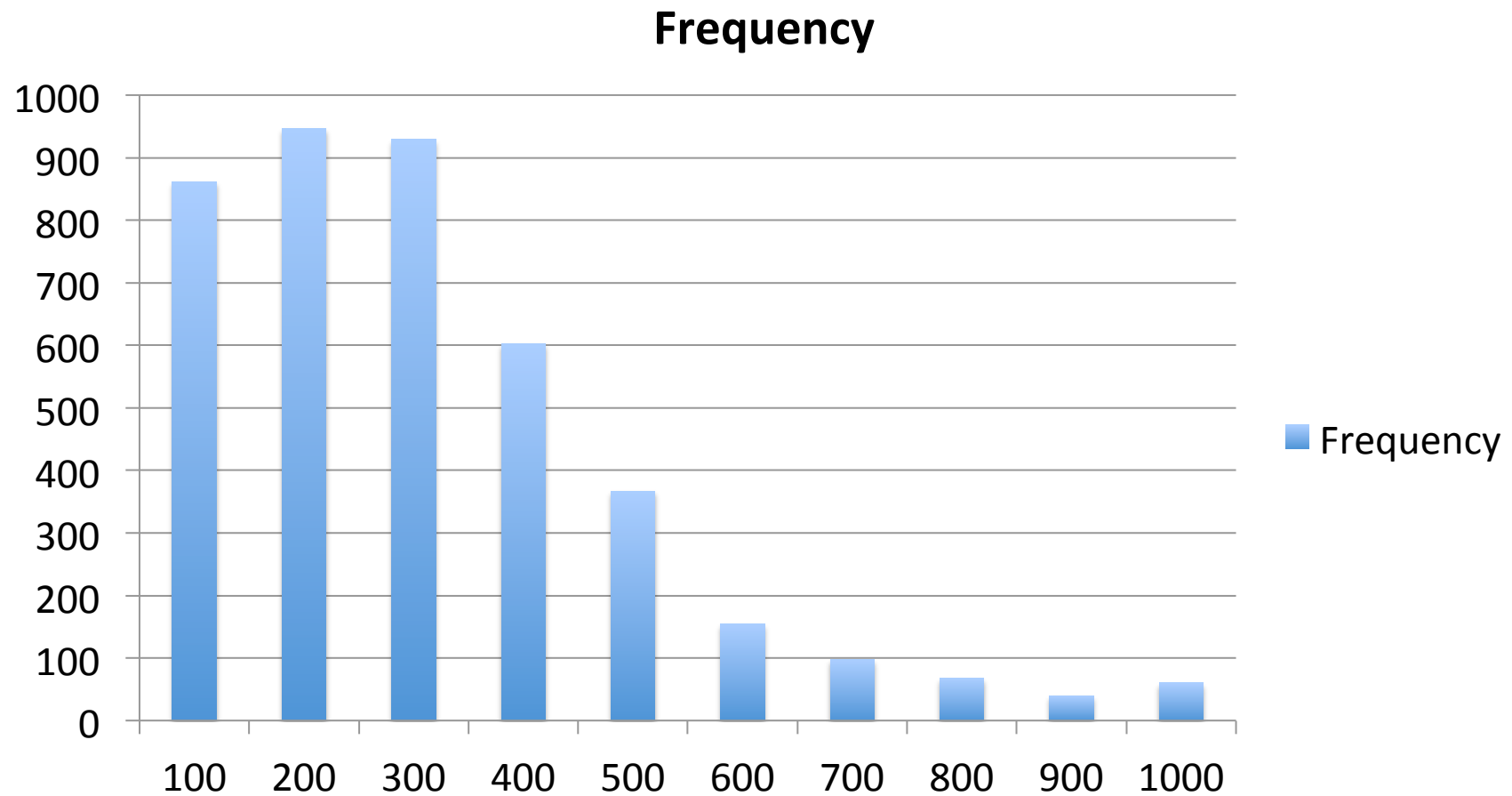
- Choosing bins according to range of values
- Assigning values to bins
- Plotting counts per bin
- x axis should be consistently scaled whenever possible

# Visualizing histograms in spreadsheets

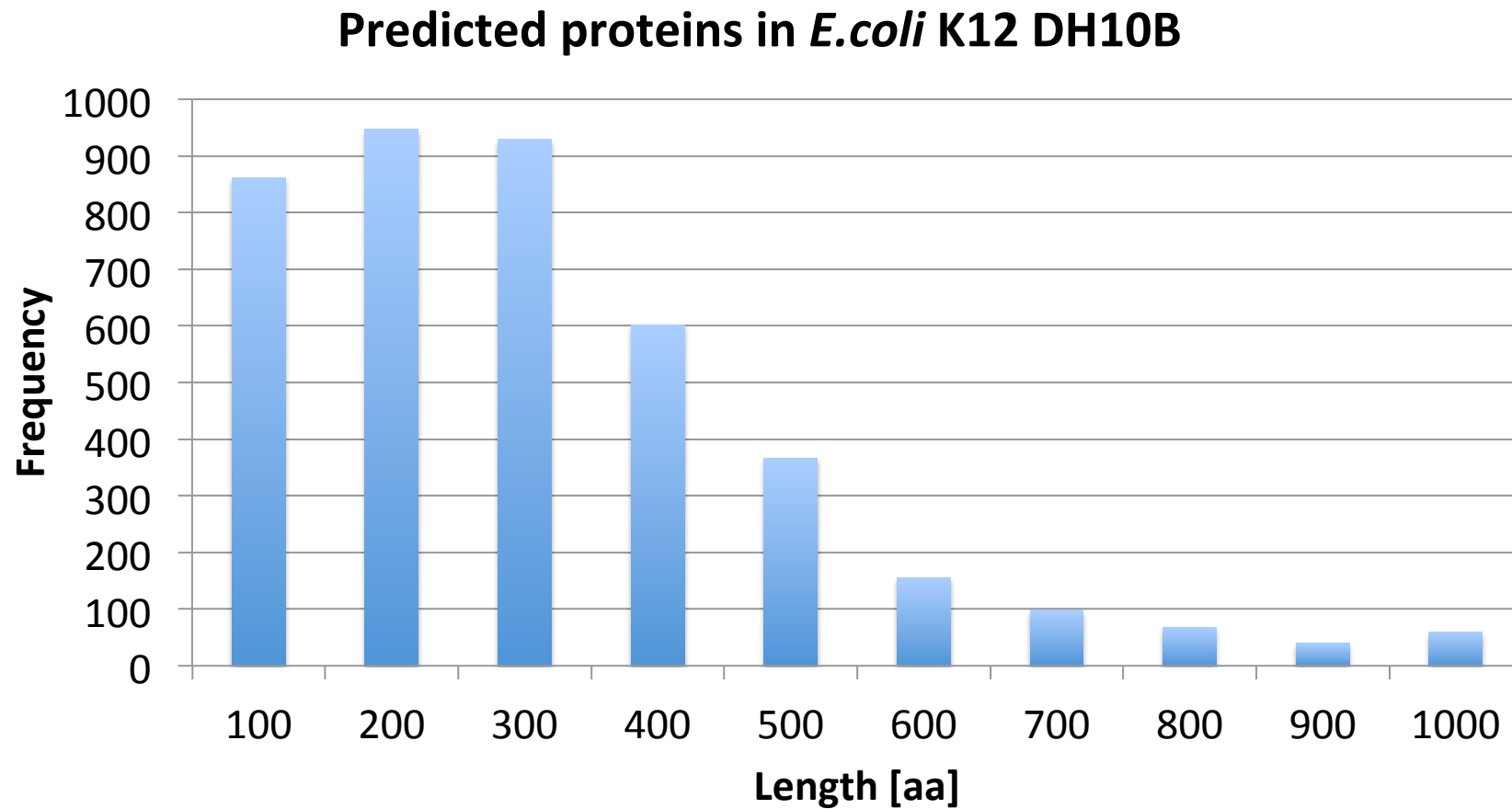
Protein length
21
820
310
428
98
258
476
317
195
188
237
161
...



# Visualizing histograms in spreadsheets



# Visualizing histograms in spreadsheets



# Visualizing histograms in spreadsheets

