WTF? Predictive Analytics
Better Data – Better Tools

Dan Pollock, Director of Data and Analytics
Dan.Pollock@deltathink.com
Data: A Neutral View
Data Tools

WTF?

What’s

The

Future
U.S. Recorded Music Revenues by Format
1973 to 2017, Format(s): All
Source: RIAA
Total # Scholarly Articles Published – Data Variance

Source: Various + Delta Think analysis.
\[ y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 \]

Where \( \beta_0, \beta_1 \) and \( \beta_2 \) are parameters to be estimated from the data. Standard practice is to find values of these parameters such that the sum of squares:

\[
\sum_{i=1}^{n} \left[ y_i - (\beta_0 + \beta_1 x_i + \beta_2 x_i^2) \right]^2
\]

is minimized. In words, we are looking for coefficients of the polynomial such that the fitted values of the polynomial are as close to the observations as possible. In matrix/vector notation what we want is the vector \( \vec{\beta} \) which satisfies:

\[
\vec{y} = X \vec{\beta}
\]

where \( \vec{\beta} = [\beta_0, \beta_1, \beta_2]^T \), \( \vec{y} = [y_1, \ldots, y_n]^T \) and

\[
X = \begin{bmatrix}
1, x_1, x_1^2 \\
1, x_2, x_2^2 \\
\vdots \\
1, x_n, x_n^2
\end{bmatrix}
\]

As we cannot invert the matrix \( X \) (it's not square for one thing), we solve the equation as follows:

\[
X^T \vec{y} = X^T X \vec{\beta}
\]

\[
(X^T X)^{-1} X^T \vec{y} = \hat{\beta}
\]
What are the implications of Plan S?
Plan S Funders’ Share of Global Expenditure on R&D

Source: R&D Magazine, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.

http://deltathink.com/sign-up-for-news-views/
Plan S Funders’ Share of Global Article Output

Source: Data provided by Dimensions from Digital Science - collected on 16 November 2018, US Science & Engineering Indicators 2016, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.
Plan S: share of scholarly articles in context

<table>
<thead>
<tr>
<th>Scholarly articles in 2017...</th>
<th>Shares of Global Research Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan S Funders Share of Global Output</td>
<td>3.5%</td>
</tr>
</tbody>
</table>

Source: Data provided by Dimensions from Digital Science - collected on 16 November 2018, US Science & Engineering Indicators 2016, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.
Plan S: share of scholarly articles in context

<table>
<thead>
<tr>
<th>Scholarly articles in 2017...</th>
<th>Shares of Global Research Articles</th>
<th>Share of Global OA Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan S Funders Share of Global Output</td>
<td>3.5%</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

Source: Data provided by Dimensions from Digital Science - collected on 16 November 2018, US Science & Engineering Indicators 2016, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.

http://deltathink.com/

sign-up-for-news-views/
Plan S: share of scholarly articles in context

<table>
<thead>
<tr>
<th>Scholarly articles in 2017...</th>
<th>Shares of Global Research Articles</th>
<th>Share of Global OA Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan S Funders Share of Global Output</td>
<td>3.5%</td>
<td>4.5%</td>
</tr>
<tr>
<td>Plan S Funders’ Share of Global Output including equivalent level of funder coverage from Germany</td>
<td>4.2%</td>
<td>5.3%</td>
</tr>
</tbody>
</table>

Source: Data provided by Dimensions from Digital Science - collected on 16 November 2018, US Science & Engineering Indicators 2016, Delta Think Open Access Data and Analytics Tool, Delta Think analysis.
Change in Market **Value** of Plan S Uptake Scenarios Compared with Current Projections

Source: Delta Think Open Access Data and Analytics Tool, Dimensions from Digital Science, Delta Think analysis.
Change in Market **Value** of Plan S Uptake Scenarios

Source: Delta Think Open Access Data and Analytics Tool, Dimensions from Digital Science, Delta Think analysis.
Change in Market **Value** of Plan S Uptake Scenarios

1. Plan S Impact - authors move to fully OA journals
2. Plan S Impact - authors move to green journals

*Source: Delta Think Open Access Data and Analytics Tool, Dimensions from Digital Science, Delta Think analysis.*
Change in Market Value of Plan S Uptake Scenarios Compared with Current Projections

Source: Delta Think Open Access Data and Analytics Tool, Dimensions from Digital Science, Delta Think analysis.
Does this data stuff really matter?
Fact 2: Our APC prices are value for money

Elsevier’s article publishing charges (APCs) are comparable with other publishers.

We publish the best quality open access articles (based on field-weighted citation impact) for both open access and hybrid titles, while our APCs remain at or below industry average.

Our APC price list is publicly available and downloadable here.

At Elsevier we do not charge subscribers for open access publications and only take into account subscription content when pricing subscription titles.

https://www.elsevier.com/about/open-science/open-access/surprising-facts
ESAC – Efficiency and Standards for Article Charges

- Collating information about agreements
- Common approaches to negotiation
- Data driven

...an example of changes in approach?
Challenges

• Common OA data needs
• Nebulous OA data sources
• Neutral data sets
• Complex calculations

OA DAT: A One-Stop Shop

• Address the challenges
• Accessible answers

Dan.Pollock  Emma.Green
@deltathink.com   @deltathink.com

http://DeltaThink.com/
sign-up-for-news-views/