

University of Vienna
VIS 2014

Assignment 1

Manfred Klaffenböck

a0301042
A033 521

The Figures

Diagram 1: Overview 2008 - 2012

(figures in millions)

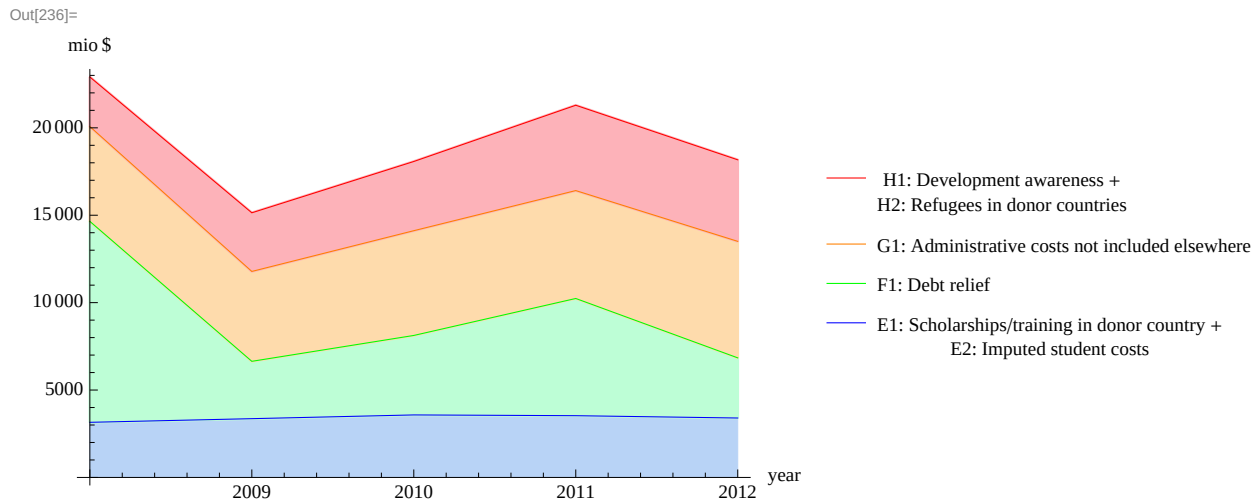


Figure 1: not transferred aids between 2008 and 2012

Diagram 2: Not transferred aids per country

(figures in millions)

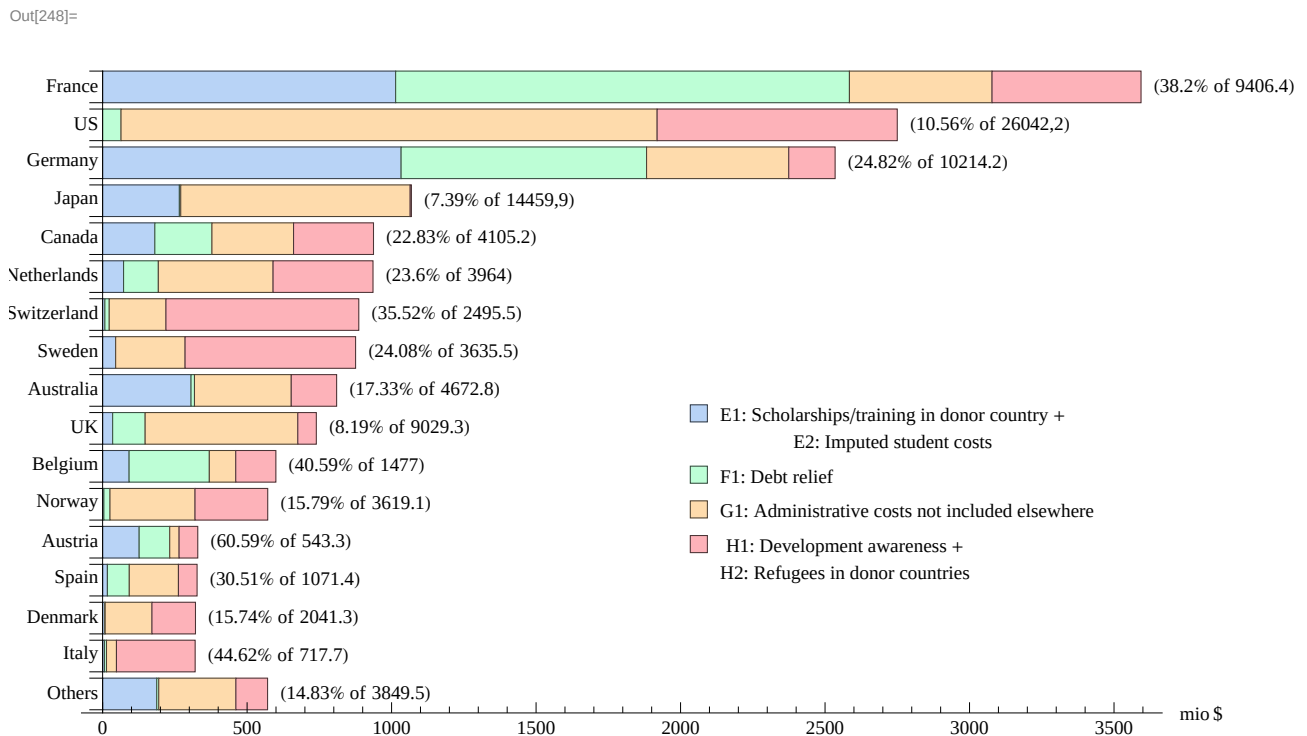


Figure 2: not transferred aids per country

Erläuterungen

Grundlegende Überlegungen

Wie im Titel der Spreadsheets bereits erkenntlich geht es um nicht bezahlte Gelder, weswegen dieser Aspekt ins Zentrum gestellt werden soll. Die Spreadsheets geben darüber hinaus weitere Infos über die Hintergründe, was mit den nicht bezahlten Geldern passiert ist bzw. wieso diese nicht ausbezahlt worden sind. Da diese Infos sowohl in "DAC countries 2012 by type of aid" als auch bei "DAC total 2008-2012 by type of aid" enthalten waren war es zudem möglich, eine Crossreferenz zwischen diesen beiden Datensätzen herzustellen.

Die absoluten nicht ausbezahlten Beträge sind in beiden Fällen mithilfe absoluter SpatialChannels repräsentiert (Fig1: Diagrammhöhe, Fig2: Balkenlänge) und lassen sich innerhalb der Diagramme leicht vergleichen (Fig1: Jahresvergleich, Fig2: Ländervergleich). Die Hintergründe wurden ebenfalls mithilfe des Spatialchannels kodiert, relativ innerhalb des Balkens. Die Farben wurden so gewählt, dass ein Vergleich zwischen den beiden Diagrammen grundsätzlich möglich ist (gleiche Farbwahl für gleiche Datensätze in den beiden Diagrammen). Für weitere Überlegungen zur Farbwahl siehe den entsprechenden Abschnitt.

Reihenfolge

Die Sortierung war im ersten Diagramm implizit durch die Jahreszahlen gegeben. Im Diagramm 2 habe ich entschieden, die Sortierung anhand der Absolutwerte der nicht ausbezahlten Gelder per Land durchzuführen. Einzige Ausnahme bildet der letzte Eintrag, wo verschiedene Länder zusammengefasst wurden. Logisch ist dieser Eintrag dennoch richtig platziert, da die darin zusammengefassten Länder alle unter Italien eingefügt werden müssten. Um zu verhindern, dass diese Reihenfolge beim Betrachter versehentlich mit "Bedeutung" oder "Ineffizienz" der jeweiligen Länder gleichgesetzt wird, habe ich zu jedem Land den Absolutwert der OECD-Gelder hinzugefügt, als auch den prozentuellen Anteil, den die nicht übermittelten Gelder in Relation zu den Gesamtgeldern ausmacht.

Gemeinsame Basis

In beiden Diagrammen wurden die Daten der Kategorie E als Basis und die Daten der Kategorie H als Spitzen gewählt, um den Zusammenhang der beiden Diagramme leichter herzustellen.

Farbwahl

Die Farben wurden so unterschiedlich gewählt, dass sie nicht versehentlich kategorisch zusammengefasst werden können (keine relativ zueinander ähnliche Farben). Desweiteren wurden sie so gewählt, dass sie auch für Personen mit einer rot-grün-Farbschwäche unterscheidbar sind (unabhängig der Tatsache, dass sowohl rot als auch grün enthalten sind).

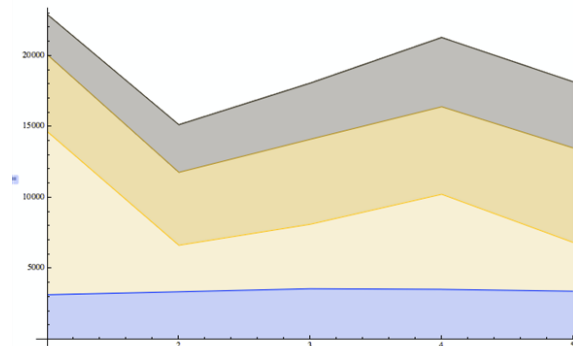


Figure 3: Diagramm 1, wie diese für Personen mit einer rot-grün-Farbschwäche erscheint

(erstellt mit Webservice unter <http://aspnetresources.com/tools/colorBlindness>)

Probleme

Designprobleme

Einige der Datensätze sind nach wie vor so gering, dass sie im Balkendiagramm untergehen (siehe z.B. Diagramm 2, F1 und H bei Japan). Mir ist leider nichts eingefallen, wie man dieses Problem umgehen könnte.

Weiters fällt auf, dass im Diagramm 1 die Legende sozusagen "auf dem Kopf" steht. Dies entspricht jedoch den relativen Positionen des Diagramms, wodurch die kognitiven Anstrengungen für das Mapping der Legende mit dem Diagramm reduziert werden. Es wäre zwar generell möglich gewesen, die Werte in umgekehrter Reihenfolge darzustellen (in Diagramm 1 rot als Basis und blau als Spitze), doch dadurch wäre die gemeinsame Basis zwischen den beiden Diagrammen zerstört worden. Somit erschien diese verkehrte Legende als das geringere Übel.

Aggregationen

Aufgrund der großen Spannweite der Daten (von Frankreich: 3593.423595 Mio. bis Island: 2.102316 Mio.) war es nicht möglich, alle

Länder im gleichen Diagramm anzuzeigen, schon gar nicht mit verschiedenen Farben innerhalb der Balkendiagramme. Aus diesem Grund wurden Ländereien unter einem gewissen Schwellenwert zu der Kategorie "Others" zusammenfasst. Da in den vorhandenen Daten ein großer Sprung zwischen Italien (320,200657 Mio) und Korea (128,409572 Mio) existiert, wurde der Schwellenwert pragmatisch zwischen diesen beiden Werten angesiedelt.

Weiters waren einige der Kategorien (vornehmlich "E2: impudent student costs" und "H01: Development Awareness") so gering, dass diese in den Diagrammen nur noch als Artefakte aufgetaucht sind. Zudem wurde durch 6 Kategorien die Farbwahl schwieriger. Da durch die kategorialen Nummern angedeutet wurde, dass eine Verbindung zwischen einigen dieser Datensätzen besteht, wurden die Kategorien E01 und E02 zu einer Farbe (blau) und die Kategorien H01 und H02 zu einer anderen Farbe (rot) zusammengefügt.

Durch diese beiden Aggregationen gingen zwar Informationen verloren (wie das bei Aggregationen eben so üblich ist), doch dies waren notwendige Maßnahmen, um die restlichen Daten adäquat darstellen zu können.

Skalierung

Ursprünglich wollte ich die Totalwerte der einzelnen Länder ebenfalls grafisch im Diagramm 2 darstellen. Doch aufgrund der wirklich großen Spannweite zwischen dem höchsten Totalwert (US: 26042,19638 Mio) und dem niedrigsten dargestellten nicht ausbezahlten Wert (Italien: 320,200657 Mio) wären die eigentlichen Daten (die nicht ausbezahlten Gelder) relativ sehr stark in den Hintergrund getreten.

Ineffizienz eines Landes

Wie bereits zuvor erwähnt, sollte durch die Reihenfolge nicht versehentlich die Ineffizienz oder "Bäuhheit" eines Landes hineininterpretiert werden. Aus diesem Grunde erschien es mir notwendig, in irgendeiner Form die relativen Anteil der nicht ausbezahlten Gelder zur Gesamtsumme im Diagramm unterzubringen. Leider ist mir hier keine grafische Lösung eingfallen, schon gar keine, die meines Wissens nach mit *Mathematica* umgesetzt werden hätte können. Zum Beispiel dachte ich daran, die Breite der Balken entsprechend den Prozentwerten zu den Gesamtleistung zu variieren (z.B. US -> 10% Balkenbreite, Österreich -> 90% Balkenbreite). Diese Lösung wäre jedoch nicht sehr intuitiv gewesen, hätte also zusätzliche Erklärung für den Betrachter erfordert, und darüber hinaus hätte ich keine Ahnung gehabt, wie dies mit *Mathematica* bewerkstelligt werden könnte.

Da ich diesen Wert trotzdem im Diagramm darstellen wollte, habe ich einen "langweiligen" Ausweg gewählt und den Prozentwert sowie den Gesamtwert eines Landes an das jeweilige Balkenende hinzugefügt.

Technische Probleme

Es war mir nicht möglich, im Diagramm 1 die Jahreszahl 2008 anzeigen zu lassen. *Mathematica* bietet zwar eine Funktion `DateListPlot`, doch diese ist sehr komplex und hat andere Nachteile, weswegen ich mich gegen diese entschieden habe.

Die Legenden haben sich dazu nur bedingt anpassen lassen. So war es nicht möglich, mehrzeilige Legenden mit einem left alignment zu versehen (wie in den Grafiken zu sehen ist, sind mehrzeilige Einträge zentriert).

Weitere Aspekte

Hintergründe Toolauswahl

Habe mich für *Mathematica* entschieden, da ich dieses Tool grundsätzlich kenne und mag. Darüber hinaus hatte ich in einem früheren Kurs schon damit gearbeitet und mir gedacht, dass ich auch über die Möglichkeiten und Limitierungen besser Bescheid weiß. Nichts desto trotz hat sich herausgestellt, dass *Mathematica* nur bedingt für komplexe Visualisierungsaufgaben geeignet ist und man durchaus einiges an Wissen mitbringen muss, um die gewünschten Effekte zu erzielen.

Es bietet jedoch mehr Flexibilität als beispielsweise Excel oder Numbers, weswegen ich mich grundsätzlich gegen diese Tools entschieden habe.

Tableau wäre mit Sicherheit eine interessante Alternative gewesen, doch ist es im Augenblick ausschließlich für Windows erhältlich und war deswegen leider keine Option für mich.