

Data Mining

Statistics from databases, visualisation and new media

14 December 2010

Stockholm University
Master Program in Statistics

<http://gauss.stat.su.se/master/statdatabaser.shtml>

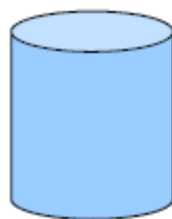
Alf Fyhrlund
Statistics Sweden - Communication/Marketing



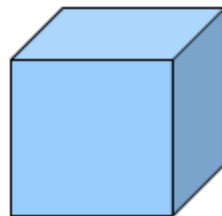
Lecture plan

- 13:00
- Data mining etc. - Alf
-
- 13:30
- Blog making and examples of visualizing data - Alf
-
- 14:00
- ----- Break 10 min.
-
- 14:10
- Showing bubbles in Trendalyzer and Motion Chart - Zhang
-
- 14:40
- How to upload your own data into Motion Chart - Zhang
-
- 15:00
- ----- Break 15min
-
- 15:15
- Group 1 - Blog making
- Group 2 - Motion Chart
-
- 16:00
- Group 1 - Motion Chart
- Group 2 - Blog making
-
- 16:45
- Integration of 2 groups
- ----- End at 17:00

Knowledge creation



Data



Statistical Analysis



Presentation



Discussion, collaboration
and validation

Knowledge



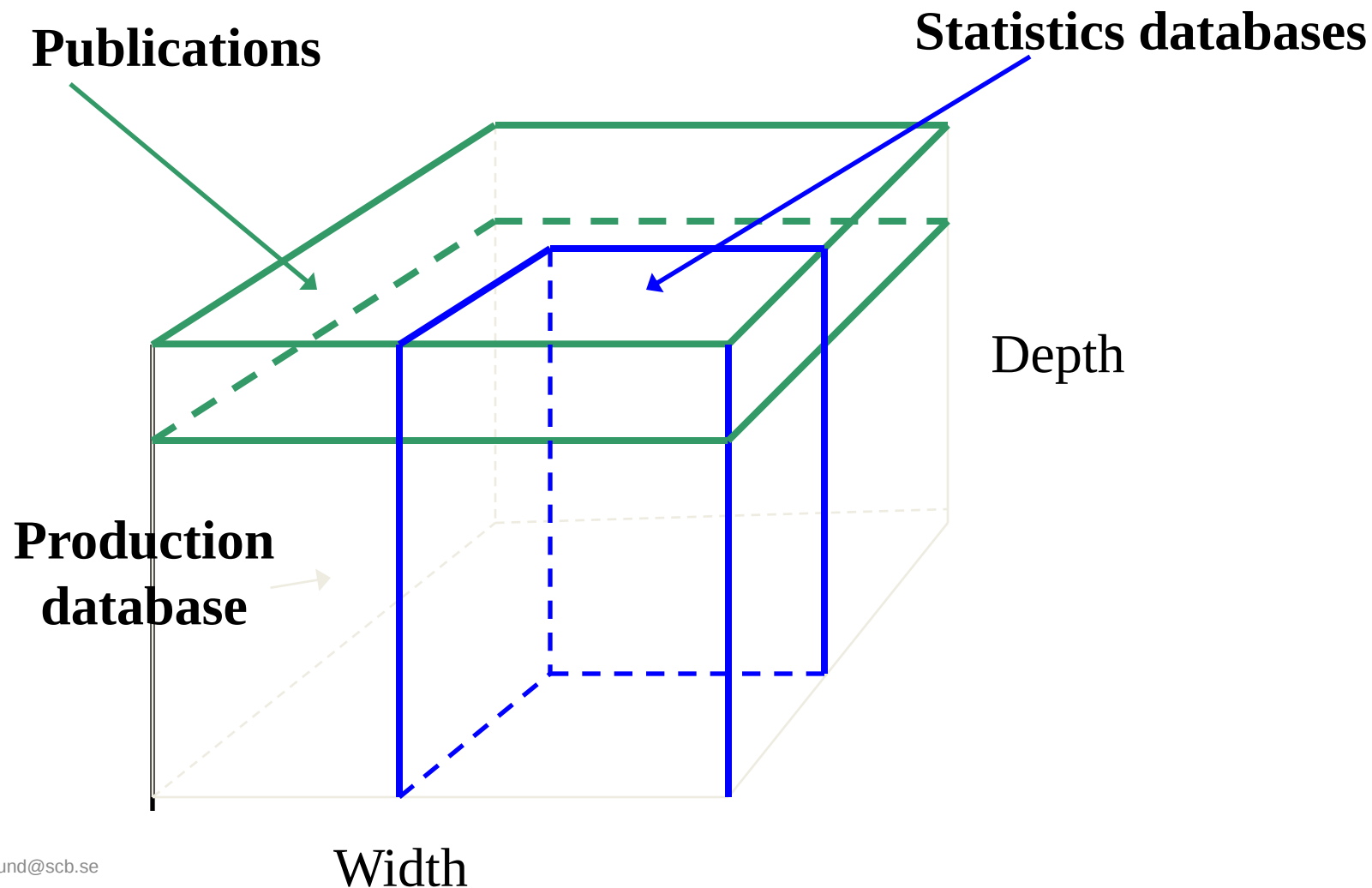
Data Mining definition

"the nontrivial extraction of implicit, previously unknown, and potentially useful information from data"

/many examples of definitions, e.g in [Wikipedia](#)

Cube of dissemination

Statistiska centralbyrån Statistics Sweden

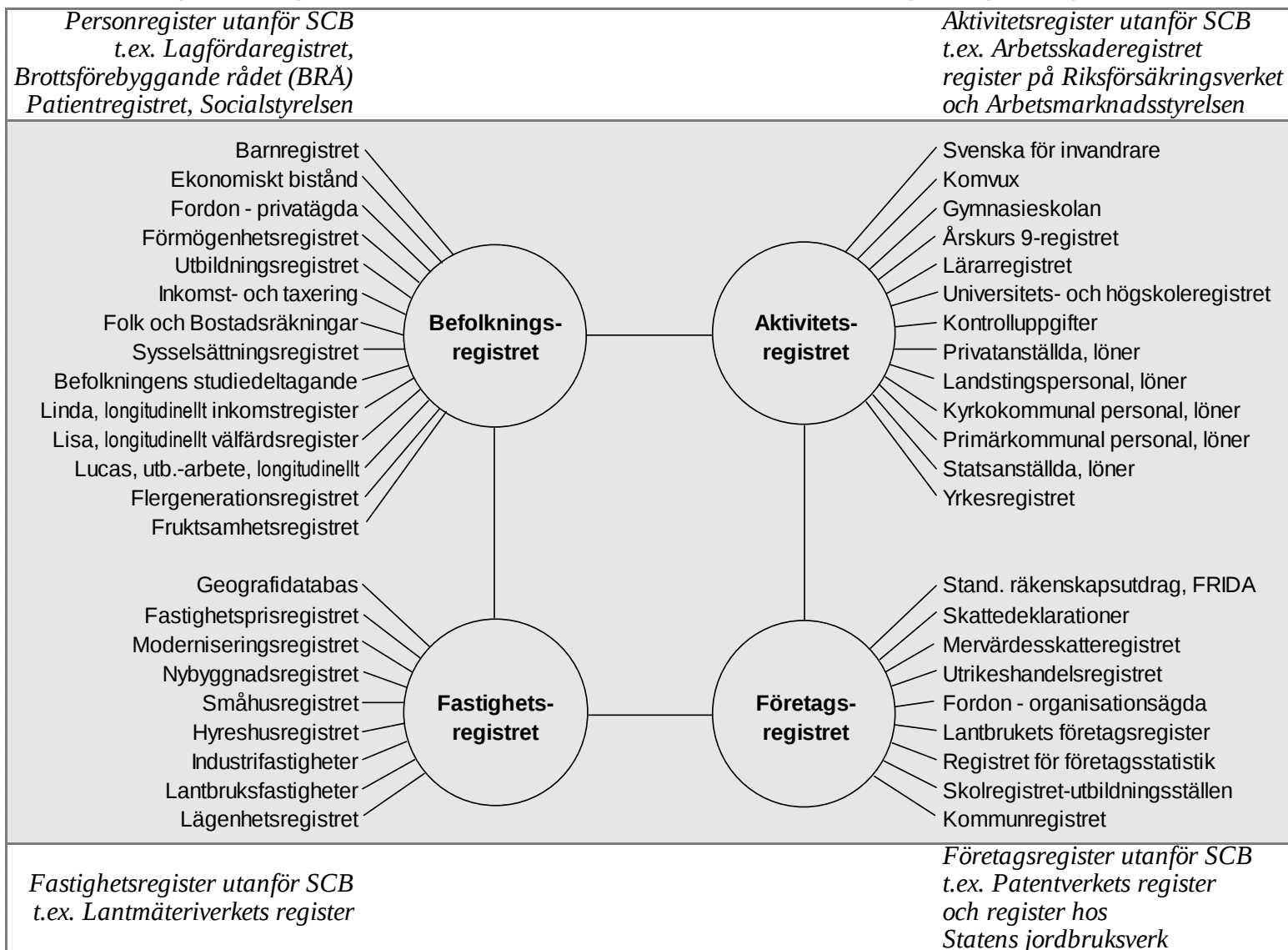


Data terminology

- **Macro data:** "Statistics"
- **Meta data:** descriptions and explanations you need to understand and use statistics correctly
- **Micro data:** individual observations which can be summarized into statistics according to user preference, under existing confidentiality

SCB: registersystem (mikrodata)

Registersystemet och olika Statistikansvariga Myndigheter



[Startsida](#)[Hitta statistik](#)[Tjänster](#)[Nyheter & press](#)[Lämna uppgifter](#)[Om SCB](#)[Kartor](#)[Dokumentationer](#)[SE ÄVEN
Vad är
Statistikdatabasen?](#)[Versionsnyheter](#)[Vanliga frågor](#)[Begreppsförklaringar](#)[Stora statistikfiler \(PC-
Axis\)](#)[Installera PC-Axis](#)[Program som kan
utnyttja data från
Statistikdatabasen](#)[Kurser](#)[Nyhetsbrev](#)

Statistikdatabasen

Nya siffror

Välj från senast uppdaterade tabeller

Välj via ämnesområde

- › Allmän statistik
- › Arbetsmarknad
- › Befolkning
- › Boende, byggnad och bebyggelse
- › Demokrati
- › Energi
- › Finansmarknad
- › Handel med varor och tjänster
- › Hushållens ekonomi
- › Hälso- och sjukvård
- › Jord- och skogsbruk, fiske
- › Kultur och fritid
- › Levnadsförhållanden
- › Miljö
- › Nationalräkenskaper
- › Näringsverksamhet
- › Offentlig ekonomi
- › Priser och konsumtion
- › Socialtjänst
- › Transporter och kommunikationer
- › Utbildning och forskning

Sök i Statistikdatabasen

[Söktips!](#)

- ☒ Sök på text
- ☐ Sök på kod (T ex 0180)

Du är inloggad med användar-id scb

Why statistics database?

- All in one place
- Structured storage
- Easy to find tables
- Flexible withdrawal options
- Edit options
- Link to metadata
- Contact person for each table



Factoring perceptions
 Globalization
 What gets measured is what gets done!
 Continuous dissemination of comparable international statistics since 1921

United Nations Statistics Division

[Home](#) |
 [Statistical Databases](#) |
 [Publications](#) |
 [Methods & Classifications](#) |
 [Meetings & Events](#) |
 [Technical Newsletters](#)

PROGRAMME OF WORK

- | [Economic Statistics](#)
- | [Demographic & Social Statistics](#)
- | [Environment & Energy Statistics](#)
- | [Development of National Statistical Systems](#)
- | [Millennium Development Goals](#)
- | [Geographic Information](#)

OUR COMMITMENTS

- | [Principles Governing International Statistical Activities](#)
- | [Fundamental Principles of Official Statistics](#)

UNSD PARTNERS

The Statistics Division is committed to the advancement of the global statistical system. We compile and disseminate global statistical information, develop standards and norms for statistical activities, and support countries' efforts to strengthen their national statistical systems. We facilitate the coordination of international statistical activities and support the functioning of the UN Statistical Commission as the apex entity of the global statistical system. For more, see [About us](#).

*Celebrating the many
contributions and
achievements of
official statistics*



20.10.2010
World Statistics Day



Service · Professionalism · Integrity

UN Statistical Commission

The **UN Statistical Commission** is the apex entity of the global statistical system. It considers and approves international statistical



Data search:

**Gateway to statistical
information from databases
of the UN and Member States**

UNSD Calendar of Events

HIGHLIGHTS, MAJOR INITIATIVES AND ACTIVITIES

[CLICK HERE](#)

IN FOCUS

[Workshop on the Methodology
and Data Compilation of
International Merchandise
Trade Statistics](#)

Phnom Penh, Cambodia,
 2-5 February 2010



Data Glossary Metadata More

Search

29 databases - 60 million records  Update calendar



Visit the Monthly Bulletin of Statistics Online

Databases

- Crime
 - International Homicide Data, UNODC
- Education
 - UIS Data Centre, UNESCO UIS 
- Energy

Updates



Country data services

-  Afghanistan
-  Albania
-  Algeria
-  Andorra
-  Angola

► Popular searches

► Feedback and reviews



**Merged Queries** ? ▾**Browse Queries** ? ▾

Frequently requested data:

☐ [Reference Series](#)☐ [Gross Domestic Product](#)**Browse Themes** ? ▾ >> Reset☐ [General Statistics](#)☐ [Agriculture and Fisheries](#)☐ [Demography and Population](#)☐ [Development](#)☐ [Economic Projections](#)☐ [Education and Training](#)☐ [Environment](#)☐ [Finance](#)☐ [Globalisation](#)☐ [Health](#)☐ [Industry and Services](#)☐ [Information and Communication
Technology](#)☐ [International Trade and Balance of
Payments](#)☐ [Labour](#)☐ [Monthly Economic Indicators](#)☐ [National Accounts](#)☐ [Prices and Purchasing Power Parities](#)☐ [Productivity](#)☐ [Public Sector, Taxation and Market
Regulation](#)☐ [Regional Statistics](#)☐ [Science, Technology and Patents](#)☐ [Social and Welfare Statistics](#)

Welcome to OECD.Stat Extracts

Four ways to access the data

OECD.Stat includes data and metadata for OECD countries and selected non-member economies.

Easy access:

- By **keyword** using "search" (e.g. [GDP](#), [Health](#), [unemployment](#), [income](#), [population](#), [labour](#), [education](#), [trade](#), [finance](#), [prices](#), [Economic Outlook](#) ...)
- By **theme** in the left-hand menu
- By accessing your **saved queries** under "browse queries"
- By **country** via the selection of key indicators in [Country statistical profiles 2009](#)

Getting Started

SELECT a dataset in the left-hand menu.

CREATE and customize your table by clicking on "current data selection".

RESHAPE your table using "pivot dimensions" to move rows and columns.

TAKE AWAY the data to Excel or CSV, print your query or save it for later use.

[more help](#)

Merged Queries ? ▾

Browse Queries ? ▾

Frequently requested data:

- ☐ Reference Series
- ☐ Gross Domestic Product

Browse Themes ? ▾

Find in Themes >> Reset

- ☐ General Statistics
- ☐ Agriculture and Fisheries
- ☐ Demography and Population
- ☐ Development
- ☐ Economic Projections
- ☐ Education and Training
- ☐ Environment
- ☐ Finance
- ☐ Globalisation
- ☐ Health
- ☐ Industry and Services
- ☐ Information and Communication Technology
- ☐ International Trade and Balance of Payments
- ☐ Labour
- ☐ Monthly Economic Indicators
- ☐ National Accounts
- ☐ Prices and Purchasing Power Parities
- ☐ Productivity
- ☐ Public Sector, Taxation and Market Regulation
- ☐ Regional Statistics
- ☐ Science, Technology and Patents
- ☐ Social and Welfare Statistics
- ☐ Non member Economies

Hourly Earnings (MEI) ⁱ

Change data selection: Subject [3 / 4] Country [37 / 39] Time & Frequency [14]



Data extracted on 03 Mar 2010 11:21 UTC (GMT) from OECD.Stat

Subject		Manufacturing, index 2005=100													
Frequency		Annual			Quarterly			Monthly							
Time		2007	2008	2009	2009			2009							2010
					Q2-2009	Q3-2009	Q4-2009	Jun-2009	Jul-2009	Aug-2009	Sep-2009	Oct-2009	Nov-2009	Dec-2009	Jan-2010
		▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼	▲▼
Country															
Australia	i	107.1	113.1	114.3	113.6	114.1	115.9	..	..	..	..	..	..	..	..
Austria	i	106.1	..	..	..	..	..	..	..	..	..	..	..	..	..
Belgium	i	104.0	106.9	109.7	109.3	110.1	110.3	..	..	..	..	..	..	..	..
Canada	i	105.1	106.9	101.6	100.8	100.7	102.0	98.7	98.4	99.8	103.9	99.8	100.6	105.6	..
Czech Republic	i	115.4	124.3	..	124.9	125.0	..	..	..	..	..	..	..	..	..
Denmark	i	107.2 (B)	110.7	113.1	114.2	113.3	113.5	..	..	113.3	..	..	113.5	..	..
Finland	i	106.3	111.4	114.0	113.6	114.1	115.3	..	..	..	..	..	..	..	..
France	i	105.7	109.0	..	111.2	111.7	..	..	..	..	..	..	..	..	..
Germany	i	103.2	106.1	..	108.1	108.2	..	..	108.2	..	..	..	..	..	..
Hungary	i	118.5	125.4	130.3	130.5	128.3	138.7	132.8	130.0	125.9	129.1	129.8	145.3	141.1	..
Ireland	i	108.3	113.5	..	115.8	..	..	..	..	..	..	..	..	..	..
Italy	i	106.2	109.8	113.4	112.9	113.4	114.6	112.9	112.9	112.9	114.2	114.5	114.5	114.8	..
Japan	i	100.9	101.0	92.9	90.9	93.3	112.0	121.6	125.8	77.5	76.7	77.3	81.8	176.9	..
Korea	i (B)	112.8	108.9	106.1	106.2	103.7	108.0	..	..	..	..	..	..	..	..
Luxembourg	i	105.7	107.0	103.1	106.9	96.5	110.4	103.5	98.6	95.6	95.4	95.1	110.4	125.8	..
Netherlands	i	103.5	107.3	110.6	110.3	111.1	111.2	110.4	111.1	111.1	111.1	111.3	111.3	111.0	111.8
New Zealand	i	108.8	113.6	..	117.2	..	..	..	..	..	..	..	..	..	..
Norway	i	110.7	117.0	..	121.2	122.3	..	..	..	..	..	..	..	..	..
Poland	i	116.0	127.3	130.5	129.5	131.6	133.3	130.4	133.4	130.5	131.0	131.3	133.4	135.1	130.0
Portugal	i	103.2	105.6	100.2	97.1	100.6	109.6	101.6	110.9	98.9	92.1	91.7	112.7	124.5	..
Slovak Republic	i	113.7	122.4	125.1	123.7	123.4	134.2	128.8	124.8	121.4	124.0	124.2	143.5	134.9	..
Spain	i	108.7	113.9	..	117.0	126.1	..	..	..	..	..	..	..	..	..
Sweden	i	106.8	111.1	..	113.9	113.0	..	113.6	113.5	112.6	113.1	113.7	114.1	..	..
Turkey	i	121.5	135.2	..	145.0	151.0	..	..	..	..	..	..	..	..	..
United Kingdom	i	108.9	112.0	..	113.2	112.2	..	112.8	112.2	111.8	112.7	113.4	114.7	..	..
United States	i	104.3	107.2	110.1	109.6	110.5	111.1	109.6	110.0	110.3	111.3	110.7	111.1	111.5	111.5
European Union	i	107.2	112.4	..	..	..	..	..	..	..	..	..	..	..	..
Euro area	i	106.1	110.5	..	..	..	..	..	..	..	..	..	..	..	..



Home

Statistics

Publications

About Eurostat

Help



Statistics Database



Release Calendars

Most popular database tables

- > GDP per capita in PPS
- > Real GDP growth rate
- > Total population
- > Unemployment rate
- > Employment rate
- > Inflation (monthly)
- > Inflation rate (annual)

Selected Statistics

- > Structural indicators
- > Euroindicators/PEEIs
- > Sustainable development indicators
- > Government finance
- > Prices (HICP)

Selected Publications

- > Eurostat Yearbook
- > European Business
- > Regional Yearbook



Combating poverty and social exclusion

Latest news releases

Press centre | RSS



€	03.03.2010	Volume of retail trade down by 0.3% in both euro area and EU27
€	02.03.2010	Industrial producer prices up by 0.7% in euro area
€	02.03.2010	Euro area inflation estimated at 0.9%
€	01.03.2010	Euro area unemployment rate at 9.9%
	01.03.2010	Organic area up by 21% in the EU between 2005 and 2008
€	26.02.2010	Euro area annual inflation up to 1.0%
€	24.02.2010	Industrial new orders up by 0.8% in euro area

Full list

Statistics in focus | Data in focus

RSS



D	03.03.2010	Harmonized indices of consumer prices – January 2010 - Issue number 5/2010
S	01.03.2010	Area under organic farming increased by 7.4% between 2007 and 2008 in the EU-27 - Issue number 10/2010
D	22.02.2010	Tourism in Europe: first results for 2009 - Issue number 4/2010
S	19.02.2010	17 % of EU citizens were at-risk-of-poverty in 2008 - Issue number 9/2010
S	09.02.2010	General economic crisis hits European port activity - Issue number 11/2010
S	04.02.2010	Pig farming in the EU, a changing sector - Issue number 8/2010
S	04.02.2010	The EU-27 construction sector: from boom to gloom - Issue number 7/2010

Full list

Search



Log in | Register | Log off



Statistics Explained



Your guide to European statistics



Country profiles



Business Cycle Clock



In the spotlight



- > Report on Greek government deficit and debt statistics

To use enhanced functionalities, please login.

New user? Please register. (Commission users don't need to register.)



[Display Custom Tree](#)
[Settings](#)
[Download Tree In PDF](#)
[Explanation](#)
[Table of Content](#)


Data Navigation Tree



Database by themes

- + General and regional statistics
- + Economy and finance
- + Population and social conditions
- + Industry, trade and services
- + Agriculture, forestry and fisheries
- + External Trade
- + Transport
- + Environment and energy
- + Science and technology



Tables by themes (new codes.xls)

- + General and regional statistics
- + Economy and finance
- + Population and social conditions
- + Industry, trade and services
- + Agriculture, forestry and fisheries
- + External Trade
- + Transport
- + Environment and energy
- + Science and technology



Tables on EU policy

- + Short-term indicators (Short-term indicators: old /new codification)
- + Structural indicators (new codes.xls)
- + Sustainable Development

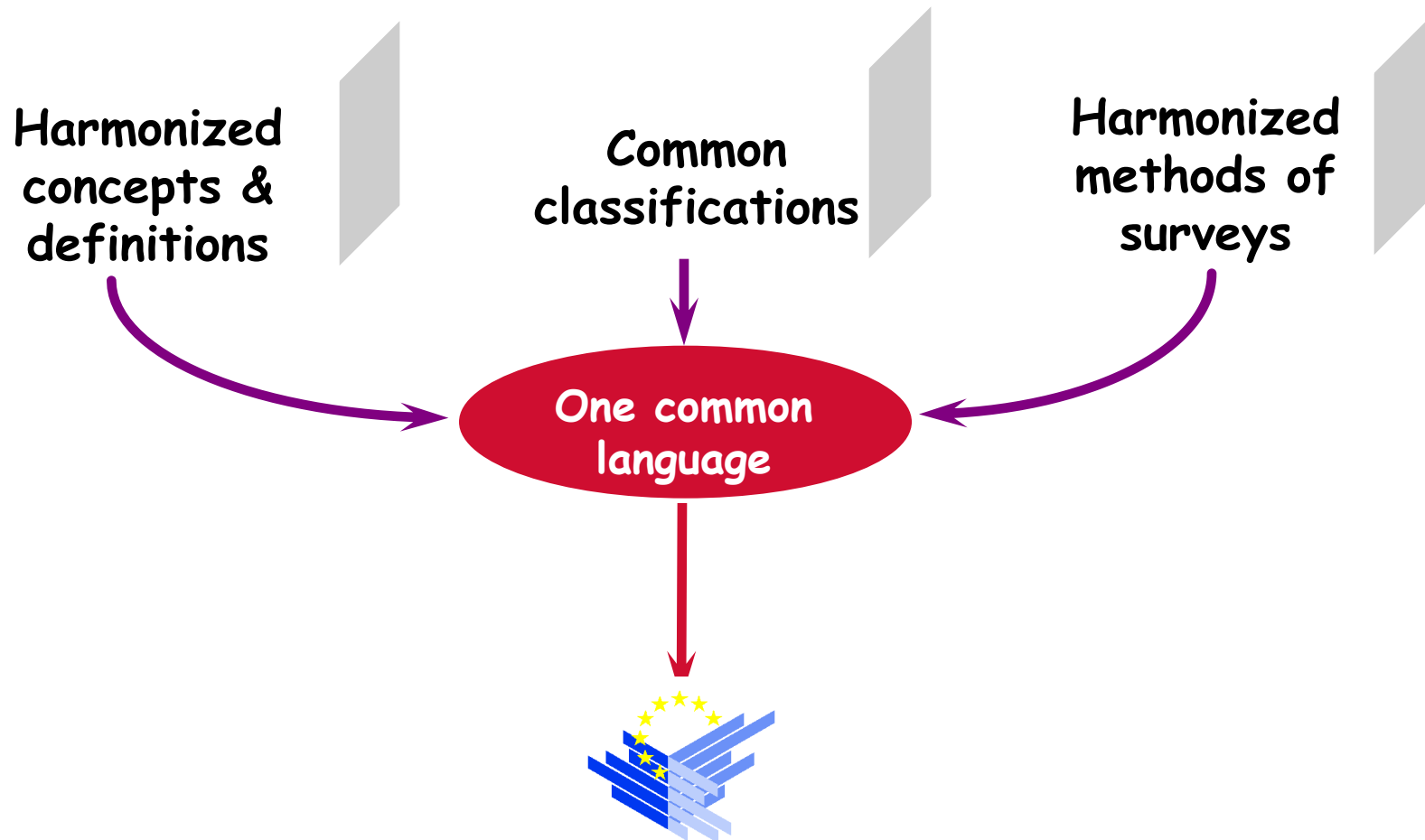


New Items (sorted by code) New



Recently Updated Items (sorted by code) Updated

To construct the European Statistical System



European Statistical System

Data Mining on the Internet

- [Statistics Sweden – Database](#)
- [United Nations – Database](#)
- [OECD Statistics – OECD.stat](#)
- [Eurostat – Database](#)
- [Open Directory Project](#)
- [Google Advanced](#)



Advanced Search

[Advanced Search Tips](#) | [About Google](#)

Use the form below and your advanced search will appear here

Find web pages that have...

all these words:

this exact wording or phrase:

[tip](#)

one or more of these words:

 OR OR

[tip](#)

But don't show pages that have...

any of these unwanted words:

[tip](#)

Need more tools?

Results per page:

Language:

File type:

Search within a site or domain:

(e.g. youtube.com, .edu)

[+ Date, usage rights, numeric range, and more](#)

Advanced Search

Topic-specific search engines from Google:

[Google Book Search](#)

[Google Code Search](#) **New!**

[Google Scholar](#)

[Google News archive search](#)

[Apple Macintosh](#)

[BSD Unix](#)

[Linux](#)

[Microsoft](#)

[U.S. Government](#)

[Universities](#)

©2010 Google

Hur överför vi statistik till kunskaper?

– Statistik metodik för detta?

Darrel Huff: Hur man ljugar med statistik

Diagram viktigt, du får inte ljuga!

Edward Tufte: The Visual Display of Quantitative Information

Ingenting får störa, ta bort allt onödigt:

"Above all else show the data!"

William Cleveland: The Elements of Graphing Data

Perception, vad uppfattar ögat? Trellis Displays (multivariat)

"Make the data stand out!"

Wallgren et al: Statistikens bilder

Bygger på Tuftes och Cleveland's principer

Datamaterialets (tabellens) struktur avgör val av diagramtyp



Datamaterialets struktur avgör diagramtyp:

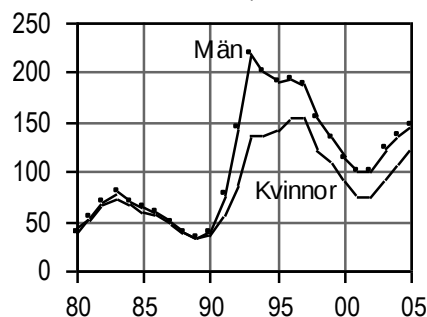
Tidsseriedata

Tvärsnittsdata

Geografiska

Bild 1. Linjediagram

Arbetslösa efter kön, 1000-tal



Envägsindelad annan tabell – del 2
län med minskning

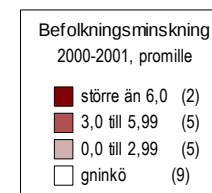
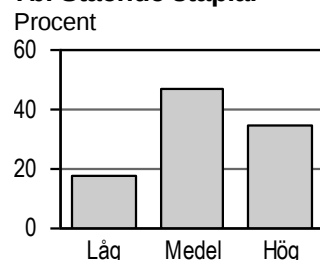
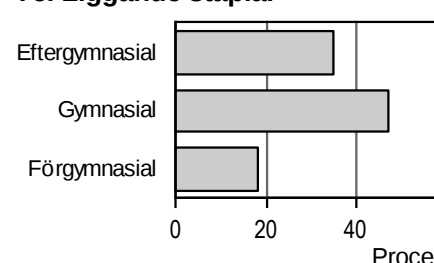


Bild 7b – 7e. Befolkningen 25–64 år efter utbildningsnivå

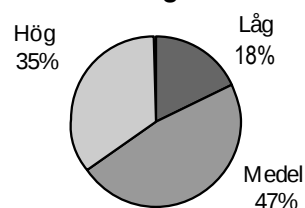
7b. Stående staplar



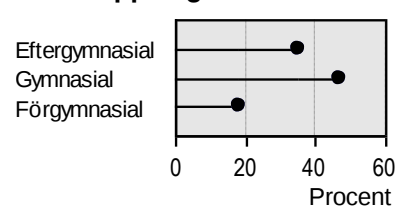
7c. Liggande staplar



7d. Cirkeldiagram



7e. Knoppdiagram



Diagrammets byggstenar: Att utforma de linjer, ytor och texter som bygger upp ett diagram



Statistics Sweden

Statistiska centralbyrån

Bild 7. Andel arbetslösa män av arbetskraften inom olika åldersklasser, 1993–2006.

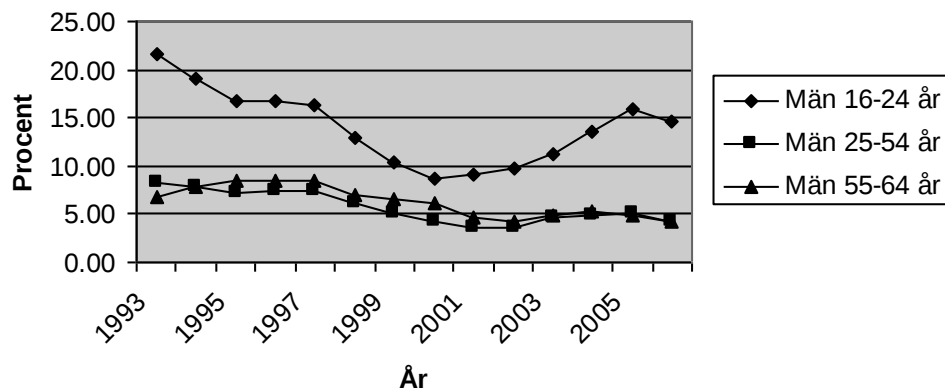


Bild 8. Andel arbetslösa män efter ålder

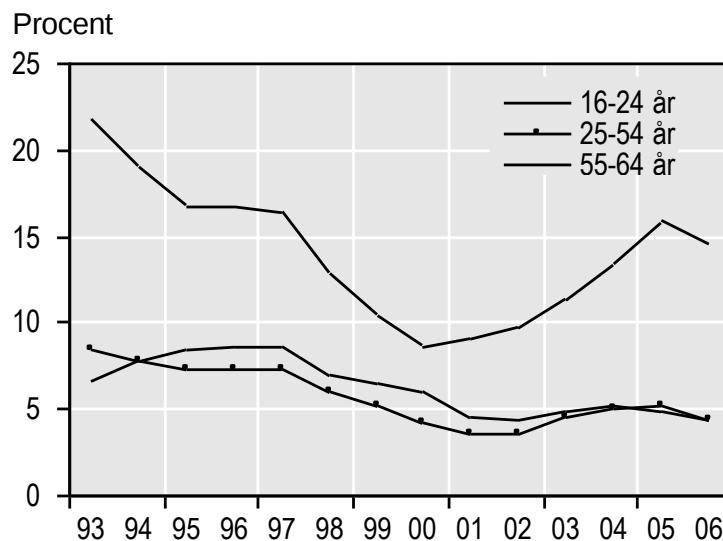


Bild 22a. Diagram i vanlig skala

Världens befolkning, miljoner

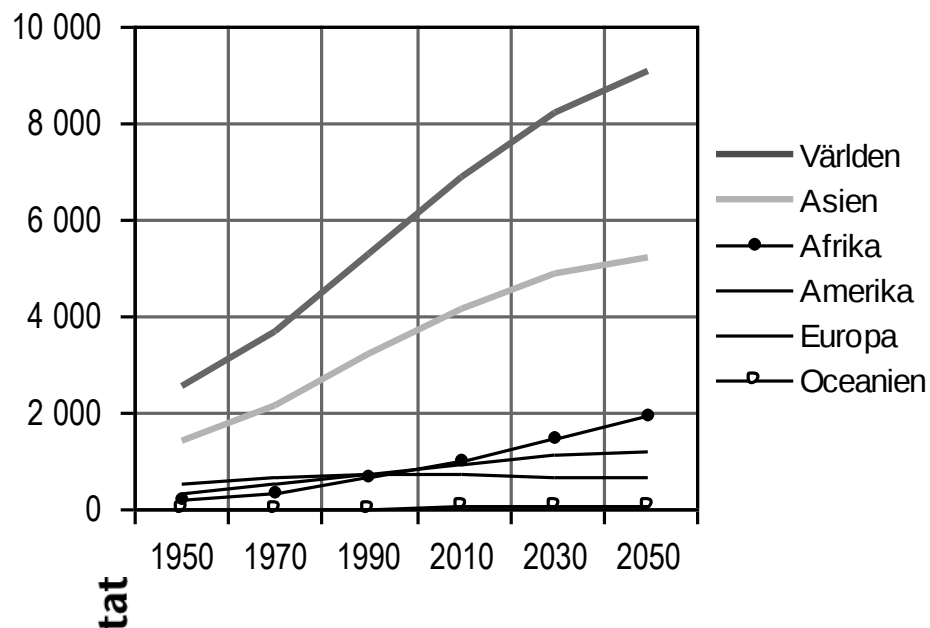


Bild 22c. Indexdiagram

Världens befolkning, index 1950=100

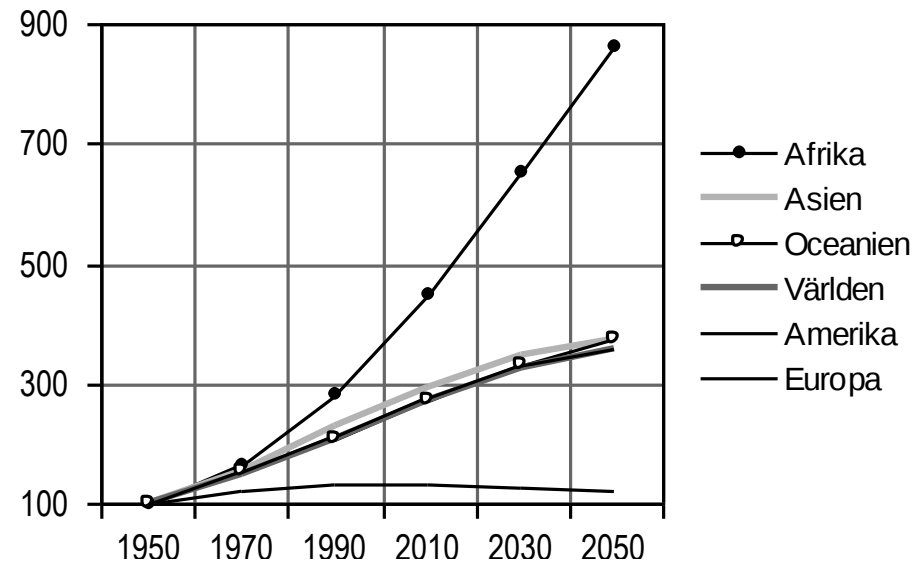
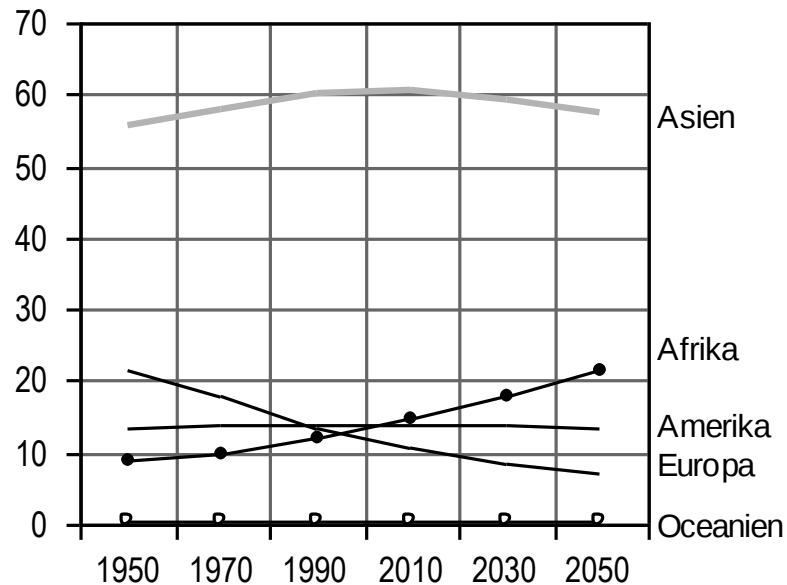


Bild 24c. Linjediagram med andelar

Andel av världens befolkning, procent



Gör flera diagram för samma datamaterial:

Det kan finnas flera mönster att illustrera

Mer och bättre analys!



Diagram utan budskap



Bild 17. Andel arbetslösa 2006 för olika åldersgrupper

Procent

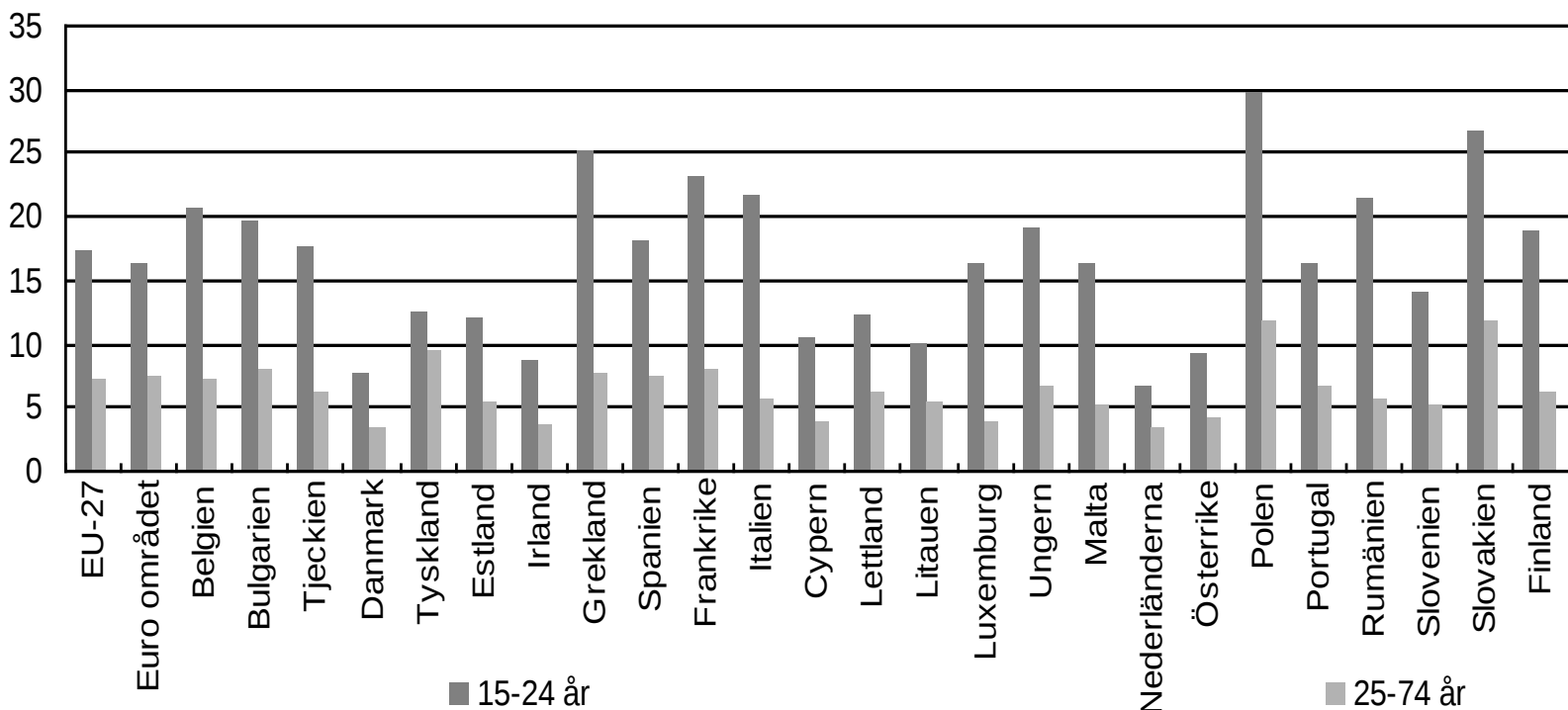


Diagram med budskap

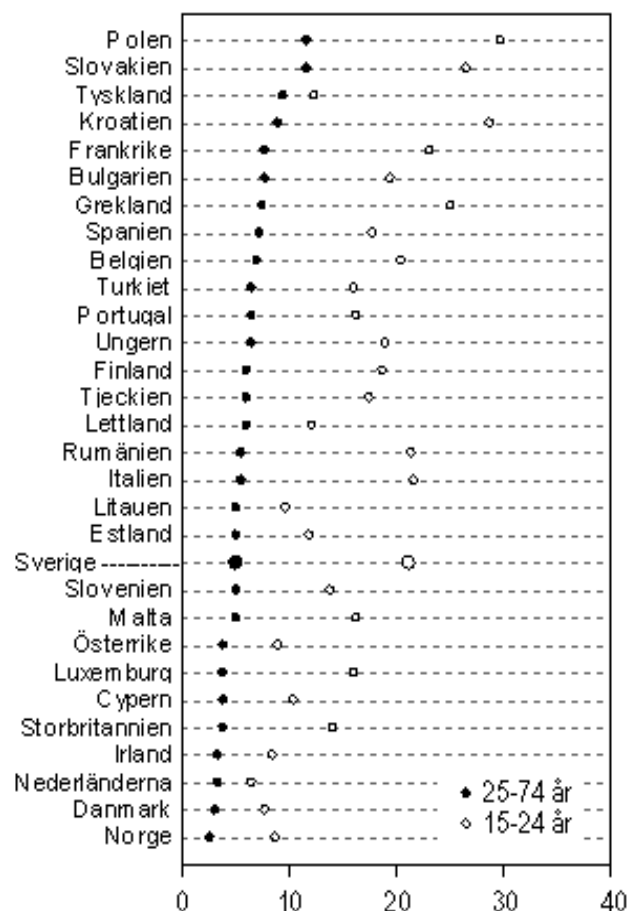


Statistics Sweden

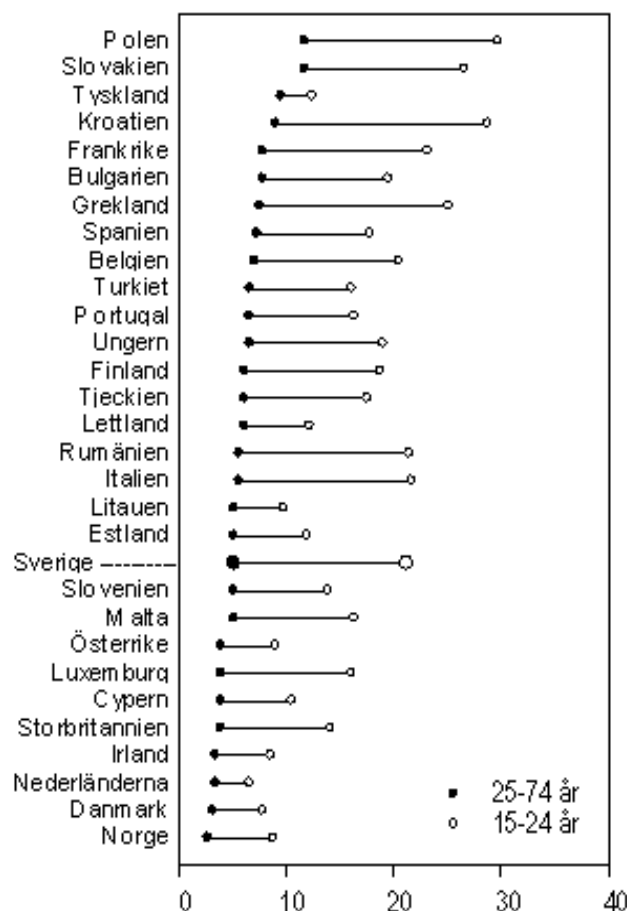
Statistiska centralbyrån

Bild 18. Varför är den svenska ungdomsarbetslösheten så hög?

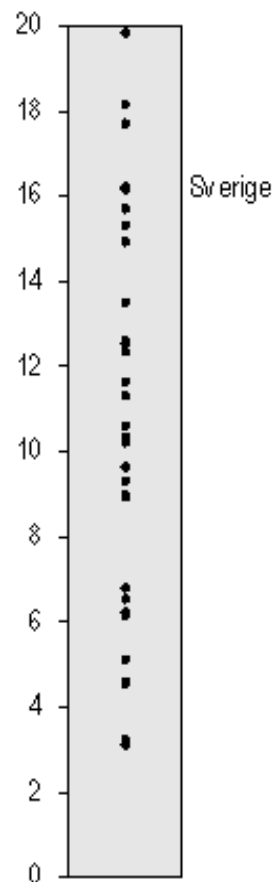
18a. Sverige: Låg vuxenarbetslöshet, men hög ungdomsarbetslöshet!
Andel arbetslösa 2006
Procent



18b. Sverige: Bland de största skillnaderna bland länderna i Europa!
Andel arbetslösa 2006
Procent



18c. Sverige på fjärde plats!
Unga - vuxna
Procentenheter



HOME
GAPMINDER WORLD
BLOG
VIDEOS
DOWNLOADS
FAQ
ABOUT

 Search

Now is the time to subscribe

Sign up for updates on new tools
and features from Gapminder.

Email:

Subscribe

NEWS!

New: Country graphs in Google search

Life expectancy, South Africa



[\(Read more\)](#)

Explore the world



Gapminder World »

See countries change in ways that challenge
old myths. You choose from the best statistics.

Explore gaps within countries:



Gaps within US

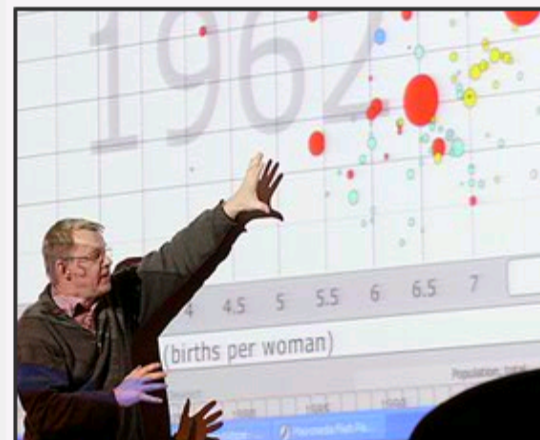
Compare US states with the countries
of the world.



Gaps within China

Compare China's provinces with the
countries of the world.

See myths get demolished



Videos »

Videos that debunk myths about Bangladesh,
CO2 emissions, urbanization and more.

Latest videos:



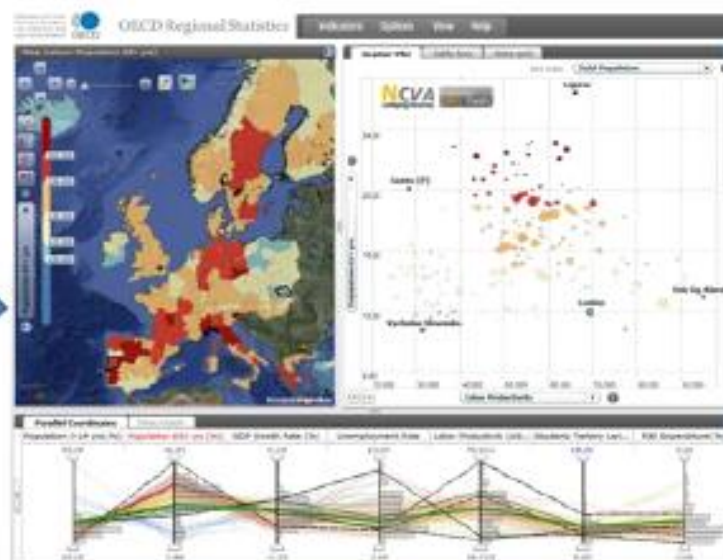
Hans Rosling: Asia's rise – how and when (TEDIndia)

Posted November 23, 2009



TED and Reddit's 10 questions to Hans Rosling

Posted September 11, 2009

NCVA
National Center for Visual Analytics**Linköpings universitet**
expanding reality

Regional statistik från Europa till en kommun i Sverige

Från statistiska rapporter till Interaktiva Dokument

PROFESSOR MIKAEL JERN NCVA

<http://ncva.itn.liu.se/>

DATA BLOG

Facts are sacred

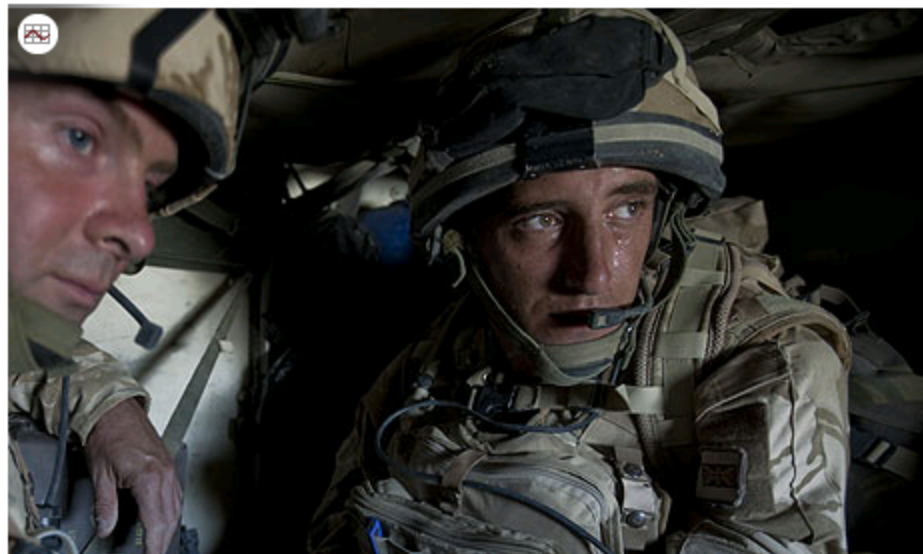


Monday 8 February 2010

British dead and wounded in Afghanistan, month by month



Posted by
[Simon Rogers](#)
15.10 GMT



B company of the Black Watch in the back of a Viking vehicle Photograph: Sean Smith

What is the human cost of the war in Afghanistan for British forces? As British troops suffer more losses in 2010, these are the latest figures - including new wounded statistics



ADOBE CREATIVE SUITE 4
DESIGN PREMIUM



INTEGRERA
InDesign och Flash
Professional

Bara med Adobe
Creative Suite 4
Design Premium

SHORTCUT TO BRILLIANT

Lär dig mer. Klicka här! [»](#)

Latest from news

Most viewed

Latest

Most commented

Last 24 hours

1. [Thousands to lose jobs as universities prepare to cope with cuts](#)
2. [Veteran congressman's death adds to Barack Obama's woes](#)
3. [Choice of 7-year-old as samba queen causes uproar in Brazil](#)
4. [Ali Dizaei, Metropolitan police commander, jailed for four years](#)
5. [Turkish girl, 16, buried alive for talking to boys](#)

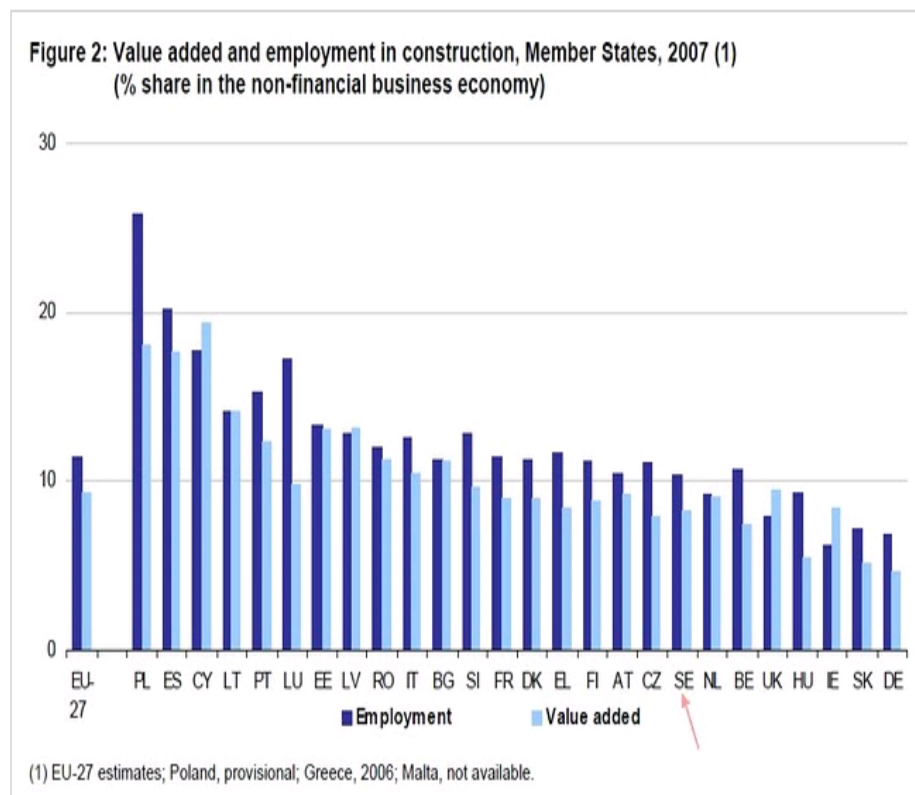
[More top stories](#)

Friday, February 5, 2010

Construction in Sweden slightly below EU-27 average

From: [The EU-27 construction sector: from boom to gloom](#)

Construction activities in the EU-27, as recorded by NACE Rev. 1.1 Section F, provided employment to an estimated 14.8 million persons in 2007 (some 11.5 % of the non-financial business economy workforce), while generating an estimated EUR 562 billion of value added (9.3 % of the non-financial business economy's total value added).



Search This Blog

powered by Google™

Europe



Click on map for some 80 000 images about Eurostat and Sweden

Latest Search Results about Eurostat and Sweden

[All Results](#) [Related Searches](#) [Wonder Wheel](#) [Time Line](#) [All Results from Swedish Sites](#) [Images](#) [Videos](#) [News](#) [Blogs](#) [Updates](#) [Books](#) [Forums](#)

Blog List

Statistics Sweden - Twitter Search
The longest winter I've witnessed. This should



[Help Center](#)[Take a Tour](#)[Docs for Work or School](#)[Team Blog](#)[Help us improve Google Docs!](#)[Documents](#)[Spreadsheets](#)[Presentations](#)

Charting: Creating a chart



Working in a spreadsheet, you can choose from five main chart types, including column, bar, and pie charts. You can also save a chart as an image and insert it into a document. To create a chart, follow these steps:

1. From the Edit page of your spreadsheet, select the cells of data you'd like to include in the chart.

Note: It helps to label the data in your spreadsheet before creating a chart. For example, if you want to chart your expenses, you might have a row of numbers labeled 'Rent' and another labeled 'Groceries.' Then you might label columns by month or week, etc. These labels will appear automatically in the window where you create and preview your chart.

Monthly Expenses	January	February
Rent	700	700
Groceries	100	30
Other	700	750

2. Select the **Pie Chart**  icon in the menu bar or choose **Insert > Chart**. The 'Create chart' window appears.
3. In the 'What type?' section, select the type of chart you'd like to create: columns, bars, lines, pie, scatter, etc.



Search Help Center

-
- ☐ Documents ☒ Spreadsheets
☐ Presentations

Known Issues with Google Docs

Review our [Known Issues page](#) for a list of known bugs and suggested workarounds.

Learn from other Google Docs users



Find answers, ask questions, and share your expertise with others in the [Google Docs Help Group](#).

Suggest a feature

Got a feature suggestion for Docs? [Let us know](#).

Google Docs Updates/Issues

3/31/2008

Hello everyone, This has been a long time coming and now it's finally here! The Google Docs team proudly presents to you Google Docs offline! Before I dive deeper, let me clarify what this really means. For Google Documents, offline viewing and ...

[Read more](#)



Updating through blog and RSS

- [Weblog – Sweden Statistics](#)
- [Google Reader](#)
- [Blog about Stats](#)

Blogging exercise

<http://maststat.blogspot.com/>

