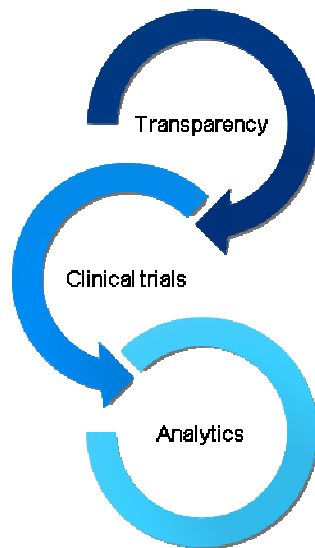


IMPROVING TRANSPARENCY OF CLINICAL DATA USING ANALYTICS

HEIDI ULZENHEIMER AND MARK LAMBRECHT, SAS



CONTENTS



- Transparency in clinical development
- Data transparency
- Analytics

TRANSPARENCY

trans·par·ent  *adjective* \tran(t)s-'per-ent\

Definition of TRANSPARENT



1 a (1) : having the property of **transmitting** light without appreciable scattering so that bodies lying beyond are seen clearly : **PELLUCID** (2) : allowing the passage of a specified form of radiation (as X-rays or ultraviolet light)

b : fine or sheer enough to be seen through : **DIAPHANOUS**

2 a : free from pretense or deceit : **FRANK**

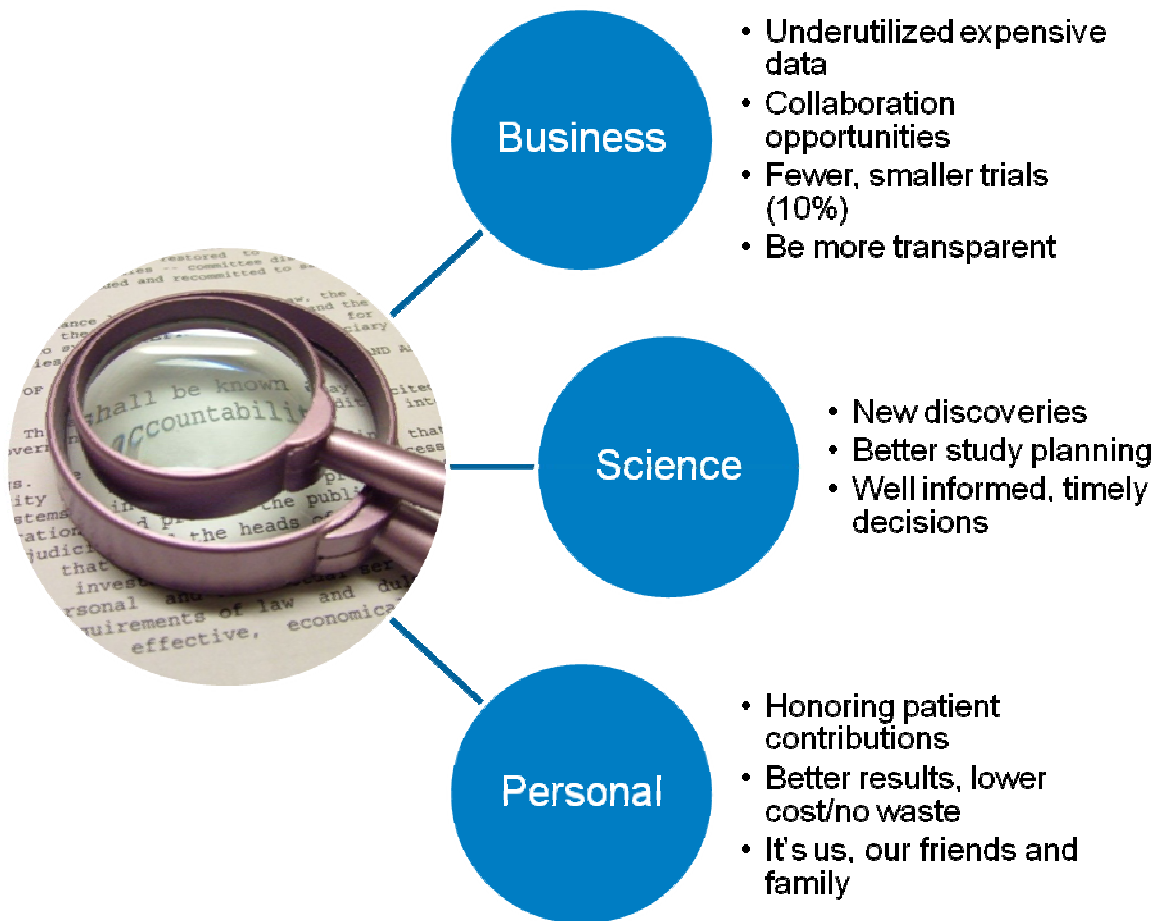
b : easily detected or seen through : **OBVIOUS**

c : readily understood

d : characterized by visibility or accessibility of information especially concerning business practices

— **trans·par·ent·ly** *adverb*

— **trans·par·ent·ness** *noun*



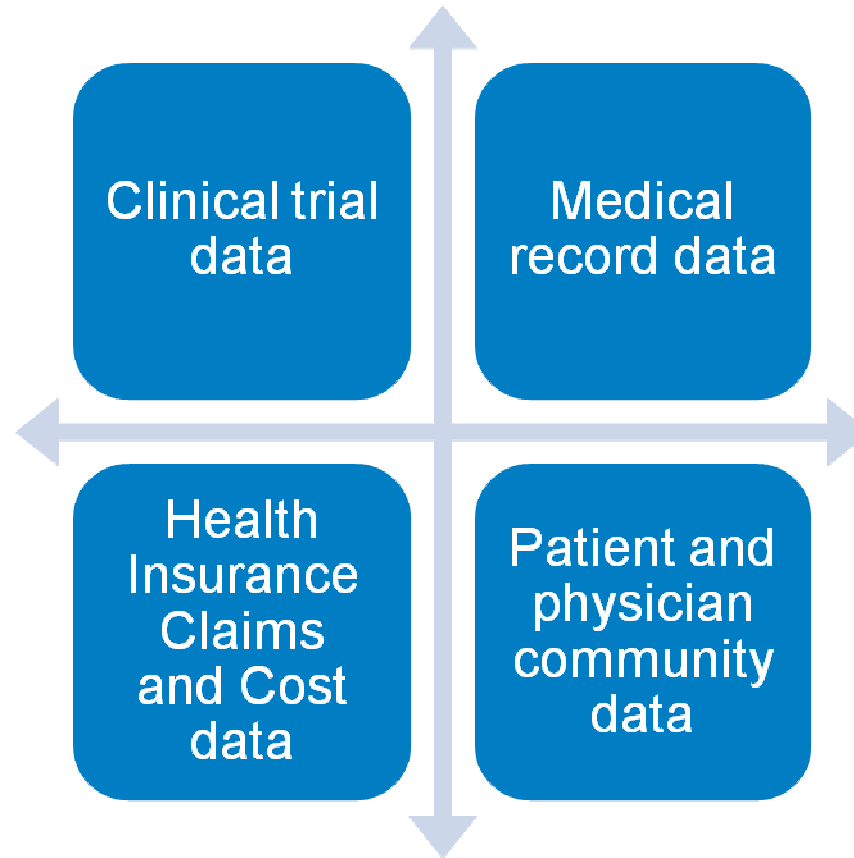
WHY NOW?



CONVERGING FORCES

- Regulatory/Health Authorities
 - EMA publish clinical-trial data and enable access to full data <https://www.clinicaltrialsregister.eu/> (ABPI monitoring) revising EU directive to include data
 - FDA <http://www.clinicaltrials.gov/> JANUS
 - NCI funded research publish data with the publication
- Patients
 - Patient advocacy groups
 - Informed and active patients
- Consortiums, other
 - CEO Roundtable LSC Task Force
 - NICE and the [AllTrials](#) campaign to make all clinical trial data available for public scrutiny

WHAT TYPE OF DATA ARE SUBJECT FOR INTERNAL AND EXTERNAL SCRUTINY ?



CONTRIBUTION OF CLINICAL DEVELOPMENT TO TRANSPARENCY

1. Enhance and analyze clinical data quality, continued dedication to audit & compliance
2. Standardize data, ready for analysis
3. Analyze data in early stage, before submission and after pooling of trial data

1. DATA QUALITY

DATA QUALITY TECHNIQUES IN CLINICAL DEVELOPMENT

- Master Data Management for clinical development

MDM FOR CLINICAL DEVELOPMENT

CONNECTING OPERATIONAL DATA SYSTEMS



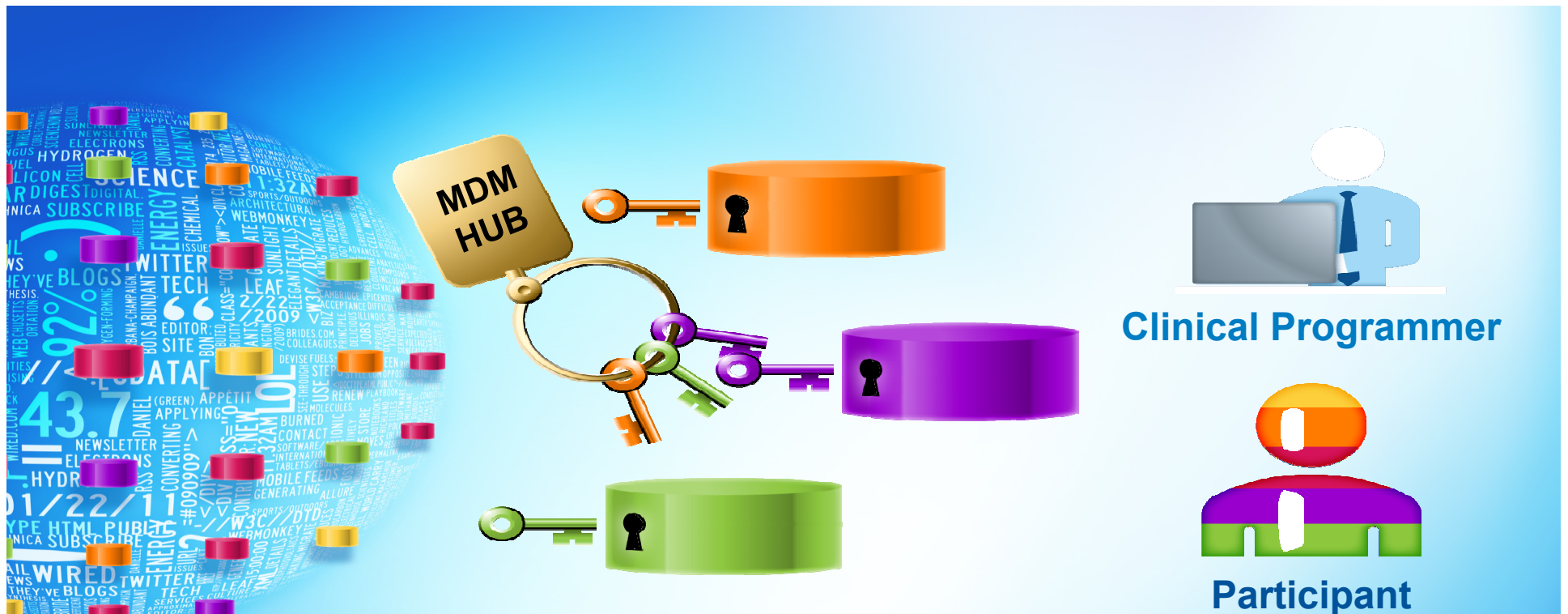
MDM FOR CLINICAL DEVELOPMENT

MANUAL CONNECTION OF PARTICIPANTS TAKES TIME



MDM FOR CLINICAL DEVELOPMENT

WHAT IF WE HAVE A MASTER DECODER KEY RING?



MDM FOR CLINICAL DEVELOPMENT

MDM IS DIFFERENT THAN DATA WAREHOUSING

	Data Warehouse	MDM
Data Content	Historical content. Multiple, disparate subject areas. Integrated data.	Single subject area. Manages and maintains links to the authoritative source.
Usage	Analytics. Reporting and provisioning of historical, multi-subject area content. User-oriented; access via SQL.	Up-to-the-moment identification. Retrieve and synchronize data from other systems. System facing; SOA access.
Integration & Load	Batch loads typical. Transactional if latency is an issue.	Integrate data upon creation. Can be batch or real-time (online).
Data Transformation	Standardization unique to each DW. Code developed/maintained for each feed.	Master data elements standardized. Data manipulation is handled by centralized rules engine. Dramatically less maintenance.
Data Remediation	Data quality checking and cleansing recommended. Infrequently occurs.	Data cleansing and validation are core. Data can be cleansed in a manner transparent to the source system.

2. DATA STANDARDIZATION

CDISC STANDARDS / SDTM WAREHOUSES

- Bringing data together across
- SAS Clinical Standards Toolkit and SAS Clinical Data Integration
- Enhanced data standardization leads to bringing together relevant domains more efficiently

QUESTIONS | ENABLED BY AN SDTM WAREHOUSE

Which studies
used efficacy
parameter xx ?

How many subjects
aged xx-xx had
event xx in studies
of a planned
duration up to xx
weeks and
containing at least
xx treated subjects ?

Were there any
trends in the
efficacy variable xx
over time when
comparing
treatment groups xx
and xx, and was
this different for
men and women ?

How many
subjects were
involved in
phase I studies ?

Which studies
had a
crossover
design (based
on information
in
title/objectives)

Questions from Vervuren, Organon

CHALLENGE

FUNDAMENTAL PARADOX OF CLINICAL DATA WAREHOUSING RESOLVED

**Scientific
flexibility**



**Well-defined
standards**

Optimized to
address scientific
and medical
objectives

Easy adaptation
to new
requirements

Optimized for
integration and
standard analysis

Easy integration
and automation



CLINICAL WAREHOUSES | REQUIREMENTS

- Provides access to organizational data
 - Different source systems in clinical : Safety, EDC, CTMS, SAP
- Data is consistent
 - Coded against MedDRA version, lab normal ranges, AECODE
- Data can be separated and merged as needed
 - User should be able to merge and use analysis-ready data as required
- Data warehouse is not just data
 - Final transformation is not end-point of data
- Separate finalized from data “in a state of flux”
 - Closed studies and ongoing studies
- Data quality is of highest importance in a warehouse
 - Don't bring data together of mixed quality

SAS CLINICAL DATA INTEGRATION



SAS CLINICAL DATA INTEGRATION FEEDS SDTM WAREHOUSE (1)

- Data Harmonization
- Data Conversion and Mapping
- Missing Data
- Possibly, Processing / Calculation (derived data)

SAS CLINICAL DATA INTEGRATION

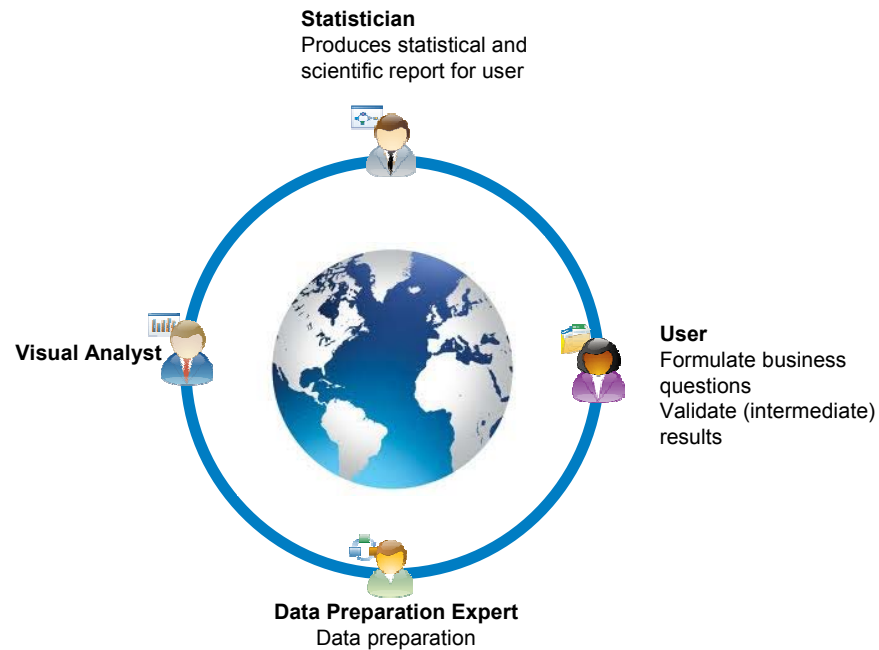


SAS CLINICAL DATA INTEGRATION FEEDS SDTM WAREHOUSE (2)

- Near real-time or high-frequent EDC extraction based on ODM
- Built-in SDTM (CDISC), ADaM, SEND, CRT-DDS, ... data models
- Check data consistency against SDTM
- Re-use for standard SDTM domains (warehouse)
- Data integration technology integrated :
 - Data Integration library for managing incremental data transfers
 - XML transfer capabilities
 - Reading from source system and writing to SAS Drug Development
- Next version of SAS CDI will be able to read define.xml and translate in standards templates

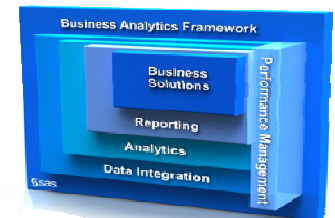
3. ANALYTICS

- Real-time and visual analysis of clinical data
- Analysis of efficacy and safety
- Cost-effectiveness or the fourth hurdle



PRODUCT PORTFOLIO OVERVIEW

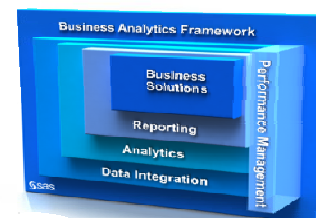
SAS® BUSINESS ANALYTICS FRAMEWORK



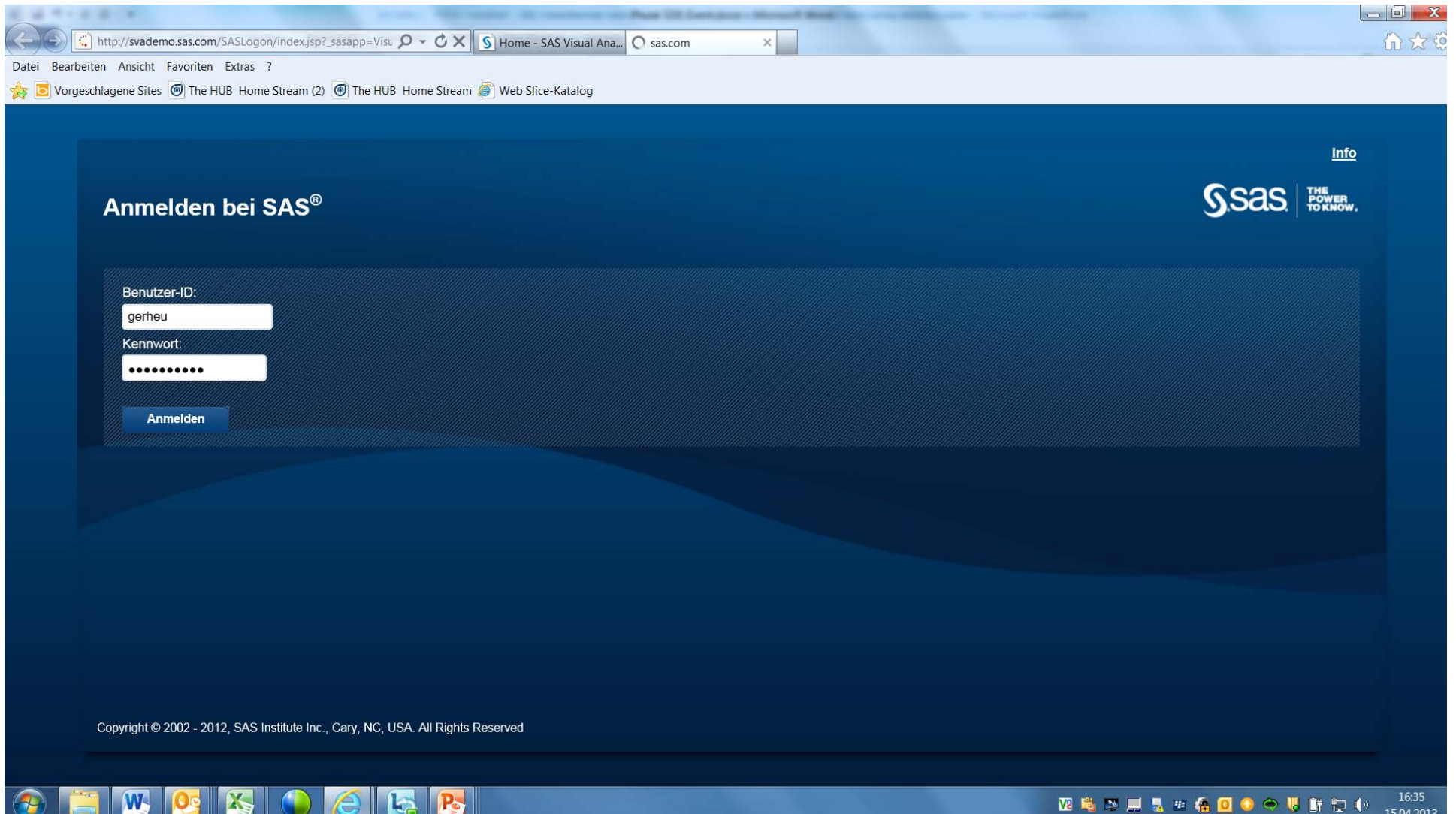
BUSINESS SOLUTIONS	SAS® Industry Solution						SAS® Regional Analytic Data Integr
	Performance Management	Customer Intelligence	Governance, Risk and Compliance	Sustainability Management	IT Intelligence	Risk Management	
REPORTING	SAS® Enterprise Business Intelligence Server						VISUAL ANALYTICS
	Web Report Studio	Add-In for MS Office	BI Dashboard	Information Delivery Portal	Enterprise Guide	Bridge to ESRI	
	AppDev Studio						
ANALYTICS	JMP® Data Discovery & Visualization						HIGH PERFORMANCE ANALYTICS
	SAS® Analytics						
	Data Mining	Text Analytics	Forecasting and Econometrics	Statistics	Operations Research	Quality Improvement	
	Model Management and Deployment						
DATA INTEGRATION	SAS® Scalable Performance Server		SAS® OLAP Server		SAS® ACCESS Data Survevor		IN DATA BASE GRID
	SAS® Enterprise Data Integration Server						
	DataFlux® Data Management						

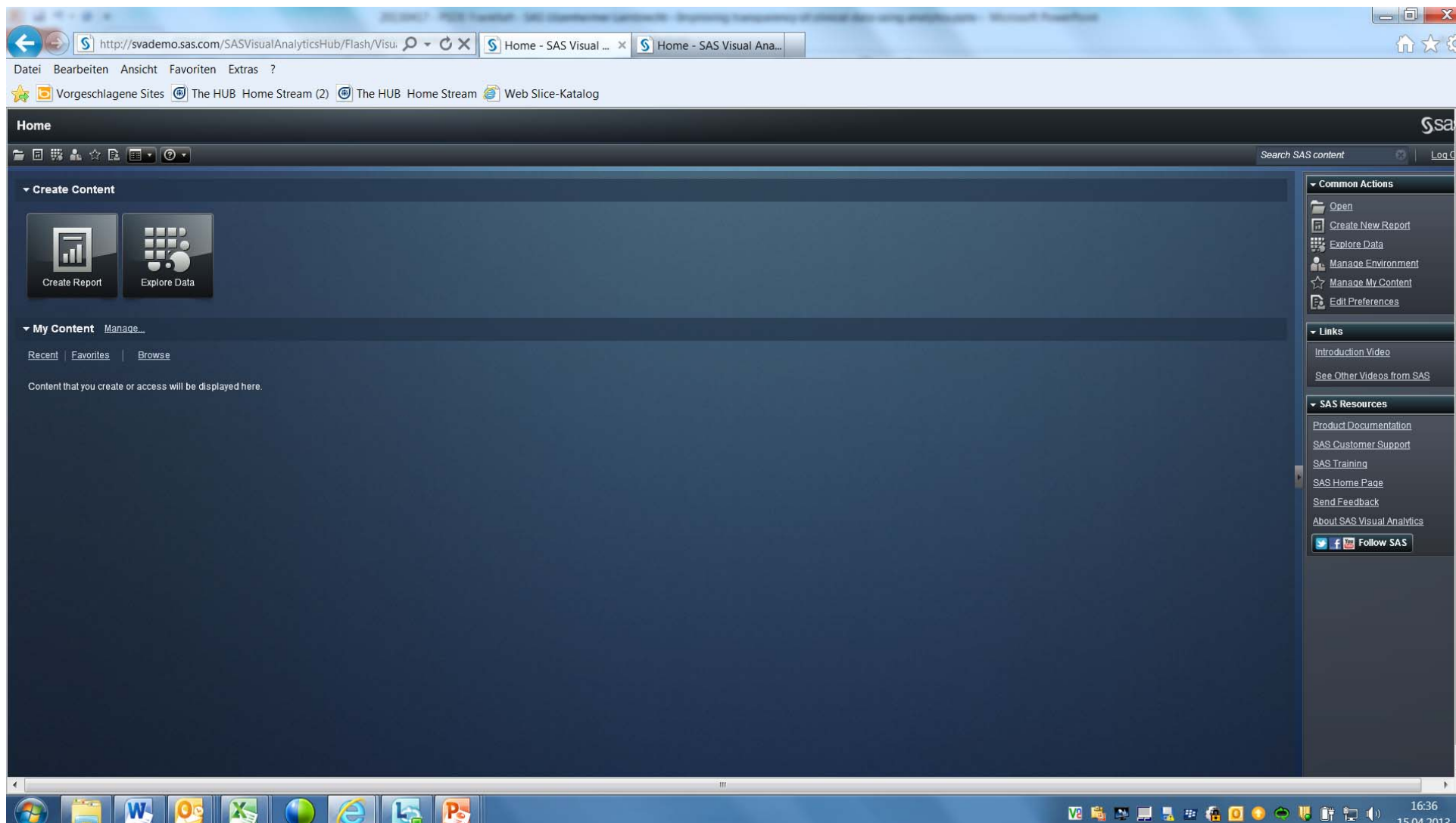
PRODUCT PORTFOLIO OVERVIEW

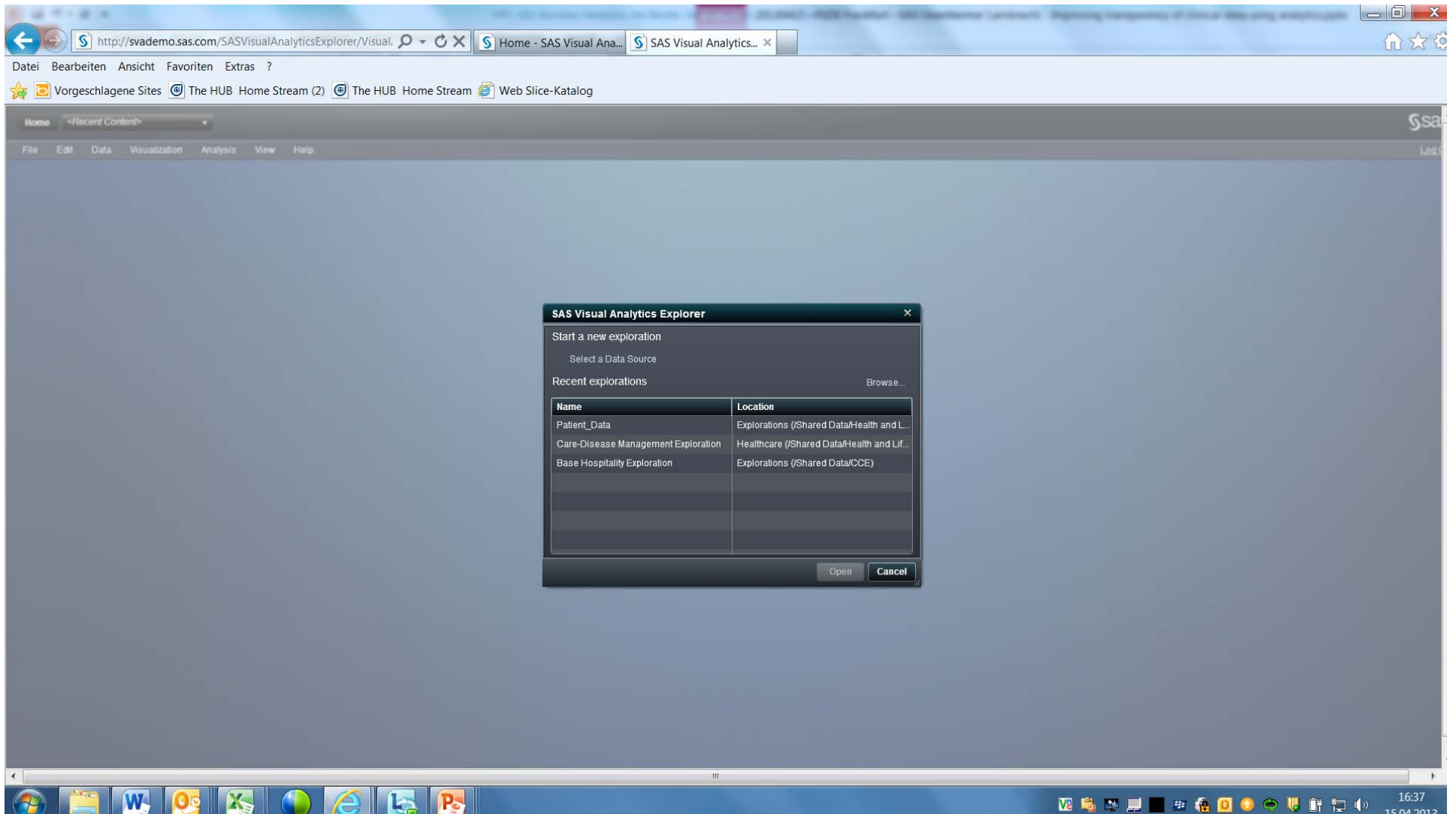
SAS® BUSINESS ANALYTICS FRAMEWORK



BUSINESS SOLUTIONS	SAS® Industry Solution						
	Performance Management	Customer Intelligence	Governance, Risk and Compliance	Sustainability Management	IT Intelligence	Risk Management	
REPORTING	SAS® Enterprise Business Intelligence Server						VISUAL ANALYTICS
	Web Report Studio	Add-In for MS Office	BI Dashboard	Information Delivery Portal	Enterprise Guide	Bridge to ESRI	
	AppDev Studio						
ANALYTICS	JMP® Data Discovery & Visualization						HIGH PERFORMANCE ANALYTICS
	SAS® Analytics						
	Data Mining	Text Analytics	Forecasting and Econometrics	Statistics	Operations Research	Quality Improvement	
	Model Management and Deployment						
DATA INTEGRATION	SAS® Scalable Performance Server		SAS® OLAP Server		SAS® ACCESS Data Survevor		IN DATA BASE GRID
	SAS® Enterprise Data Integration Server						
	DataFlux® Data Management						







Home - SAS Visual Ana...Patient_Data - SAS ...

Datei Bearbeiten Ansicht Favoriten Extras ?

Vorgeschlagene SitesThe HUB Home Stream (2)The HUB Home StreamWeb Slice-Katalog

HomePatient_Data

FileEditDataVisualizationAnalysisViewHelp

PATIENT_DATA

AE_STARTDT1,104

AE_STOPDT1,113

AE110

AE210

AE310

AGE_RANGE7

City5,000+

Country203

DATE1,095

DATE_MONTH_NAME12

DATE_YEAR4

EFFICACY4

GENDER2

Geography

PRIMARY_MED8

PRIME_DOSAGE_FORM2

SECONDARY_DOSAGE_...1

SECONDARY_MED6

SEVERITY3

State52

TYPE4

WEEKDAY_NAME7

25-Hydroxyvitamin D ng/ml

ABDOMINAL PAIN

Acetoacetate mg/ml

AE_DURATION

ae1 Blank

AGE

Alanine (ALT) units/L

Alcohol mg/dl

PropertyValue

Name

Role

Model type

Format

Aggregation

Expression

Drop a data item here

FiltersComments

Global Filters

Drop data items here to create global filters.

Geo Map

Bubble Plot

Heat Map

Correlation Detail

Box Plot

16:3715.04.2013

Browser window showing the SAS Visual Analytics Explorer interface. The URL is <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual>. The page title is "Patient_Data - SAS ...".

The interface includes a menu bar (Datei, Bearbeiten, Ansicht, Favoriten, Extras, ?) and a toolbar with icons for file operations, data manipulation, and visualization. The main workspace is a large blue area with a dashed border and a central icon indicating "Drop a data item here".

The left sidebar displays a list of data items under the "PATIENT_DATA" dataset. A context menu is open over the "PATIENT_DATA" dataset, showing options: Refresh Data, Change Data Source..., New Hierarchy..., New Calculated Item..., Edit Selected Item..., Category, Measure, Geography, Data Properties..., Measure Details..., and Data Source Details....

The bottom of the sidebar shows a table with columns "Property" and "Value".

The right sidebar contains a "Filters" tab and a "Comments" tab. The "Filters" tab shows a section for "Global Filters" with the instruction "Drop data items here to create global filters".

The bottom status bar shows the system clock: 16:41, 15.04.2013.

Property	Value
Name	
Role	
Model type	
Format	
Aggregation	
Expression	

Browser window showing the SAS Visual Analytics Explorer interface. The URL is <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual...>. The interface includes a menu bar (Datei, Bearbeiten, Ansicht, Favoriten, Extras, ?) and a toolbar. The main area displays a list of data items under the 'PATIENT_DATA' dataset, including AE_STARTDT, AE_STOPDT, AE1, AE2, AE3, AGE_RANGE, City, Country, DATE, DATE_MONTH_NAME, DATE_YEAR, EFFICACY, GENDER, Geography, PRIMARY_MED, PRIME_DOSAGE_FORM, SECONDARY_DOSAGE, SECONDARY_MED, SEVERITY, State, TYPE, WEEKDAY_NAME, 25-Hydroxyvitamin D ng/ml, ABDOMINAL_PAIN, Acetaminophen mg/ml, AE_DURATION, ae1 Blank, AGE, Alanine (ALT) unit/L, and Alcohol mg/dl. A central dialog box titled 'USER.DEMOS.PATIENT_DATA' displays the message 'Data was not filtered during import' and provides summary statistics: Total rows: Calculate, Returned rows: 27,575,488, Columns shown: 131 of 131, and Filtered: False. The dialog box has a 'Close' button. The bottom status bar shows the time 16:41 and date 15.04.2013.

Browser window showing the SAS Visual Analytics Explorer interface. The URL is <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual...>. The interface includes a menu bar (Datei, Bearbeiten, Ansicht, Favoriten, Extras, ?) and a toolbar. The main area displays a list of data items under the 'PATIENT_DATA' dataset, including AE_STARTDT, AE_STOPDT, AE1, AE2, AE3, AGE_RANGE, City, Country, DATE, DATE_MONTH_NAME, DATE_YEAR, EFFICACY, GENDER, Geography, PRIMARY_MED, PRIME_DOSAGE_FORM, SECONDARY_DOSAGE, SECONDARY_MED, SEVERITY, State, TYPE, WEEKDAY_NAME, 25-Hydroxyvitamin D ng/ml, ABDOMINAL_PAIN, Acetaminophen mg/ml, AE_DURATION, ae1 Blank, AGE, Alanine (ALT) unit/L, and Alcohol mg/dl. A central dialog box titled 'USER.DEMOS.PATIENT_DATA' displays the message 'Data was not filtered during import' and provides summary statistics: Total rows: Calculate, Returned rows: 27,575,488, Columns shown: 131 of 131, and Filtered: False. The dialog box has a 'Close' button. The bottom status bar shows the time 16:41 and date 15.04.2013.

Browser window showing the SAS Visual Analytics Explorer interface. The URL is <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual>. The interface includes a menu bar (Datei, Bearbeiten, Ansicht, Favoriten, Extras, ?), a toolbar, and a sidebar with a list of data items (PATIENT_DATA, AE_STARTDT, AE_STOPDT, AE1, AE2, AE3, AGE_RANGE, City, Country, DATE, DATE_MONTH, DATE_YEAR, EFFICACY, GENDER, Geography, PRIMARY_MED, PRIME_DOSAGE_FORM, SECONDARY_DOSAGE, SECONDARY_MED, SEVERITY, State, TYPE, WEEKDAY_NAME, 25-Hydroxyvitamin D ng/ml, ABDOMINAL PAIN, Acetoacetate mg/ml, AE_DURATION, ae1 Blank, AGE, Alanine (ALT) units/L, Alcohol mg/dl).

The main workspace displays a large blue area with the text "Drop a data item here". A context menu is open over the workspace, showing options: Refresh Data, Change Data Source..., New Hierarchy..., New Calculated Item..., Edit Selected Item..., Category, Measure, Geography, Data Properties..., Measure Details..., and Data Source Details....

On the right side, there is a "Filters" panel with a "Global Filters" section and a "Comments" section. The "Global Filters" section contains the text: "Drop data items here to create global filters."

The bottom of the interface shows a Windows taskbar with various application icons (Internet Explorer, Word, Excel, etc.) and a system tray with the time 16:43 and date 15.04.2013. The SAS logo and "POWER TO KNOW" tagline are visible in the bottom right corner.

Property	Value
Name	
Role	
Model type	
Format	
Aggregation	
Expression	

Home Patient_Data

File Edit Data Visualization Analysis View Help

Data

PATIENT_DATA

AE_STARTDT 1,104

AE_STOPDT 1,113

AE1 10

AE2 10

AE3 10

AGE_RANGE 7

City 5,000+

Country 203

DATE 1,095

DATE_MONTH_NAME 12

DATE_YEAR 4

EFFICACY 4

GENDER 2

Geography

PRIMARY_MED 8

PRIME_DOSAGE_FORM 2

SECONDARY_DOSAGE_ 1

SECONDARY_MED 6

SEVERITY 3

State 52

TYPE 4

WEEKDAY_NAME 7

25-Hydroxyvitamin D ng/ml

ABDOMINAL PAIN

Acetoacetate mg/ml

AE_DURATION

ae1 Blank

AGE

Alanine (ALT) units/L

Alcohol mg/dl

Filters Comments

Global Filters

Drop data items here filters

Data Properties

Name	Data Type	Role	Model Type	Format	Aggregation
25-Hydroxyvitamin D ng/ml	Number	Measure	Continuous	Float	Sum
ABDOMINAL PAIN	Number	Measure	Continuous	Numeric	Sum
Acetoacetate mg/ml	Number	Measure	Continuous	Float	Sum
AE1	Text	Category	Discrete	\$	None
ae1 Blank	Number	Measure	Continuous	Numeric	Sum
AE2	Text	Category	Discrete	\$	None
AE3	Text	Category	Discrete	\$	None
AE_DURATION	Number	Measure	Continuous	Numeric	Sum
AE_STARTDT	Date	Category	Continuous	Month, Day, Year	None
AE_STOPDT	Date	Category	Continuous	Month, Day, Year	None
AGE	Number	Measure	Continuous	Numeric	Sum
AGE_RANGE	Text	Category	Discrete	\$	None
Alanine (ALT) units/L	Number	Measure	Continuous	Float	Sum
Alcohol mg/dl	Number	Measure	Continuous	Float	Sum

OK Cancel

Browser address bar: <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual...>

Browser tabs: SAS Logon Manager, Patient_Data - SAS ...

Page header: Home Patient_Data

Menu: File Edit Data Visualization Analysis View Help

Left Panel: PATIENT_DATA

Property	Value
Name	AE2
Role	Category
Model type	Discrete
Format	\$
Aggregation	None

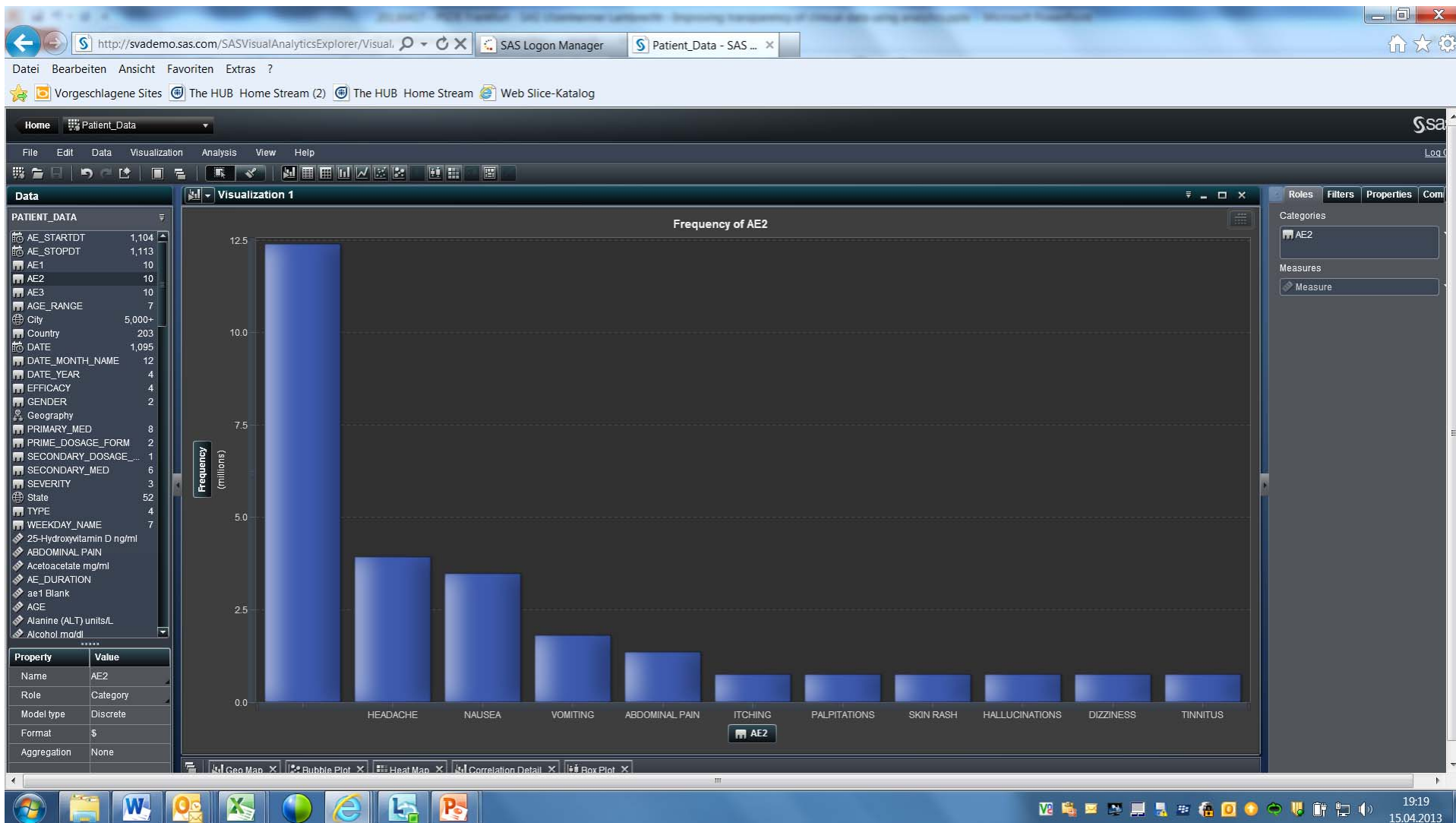
Central Canvas: Drop a data item here

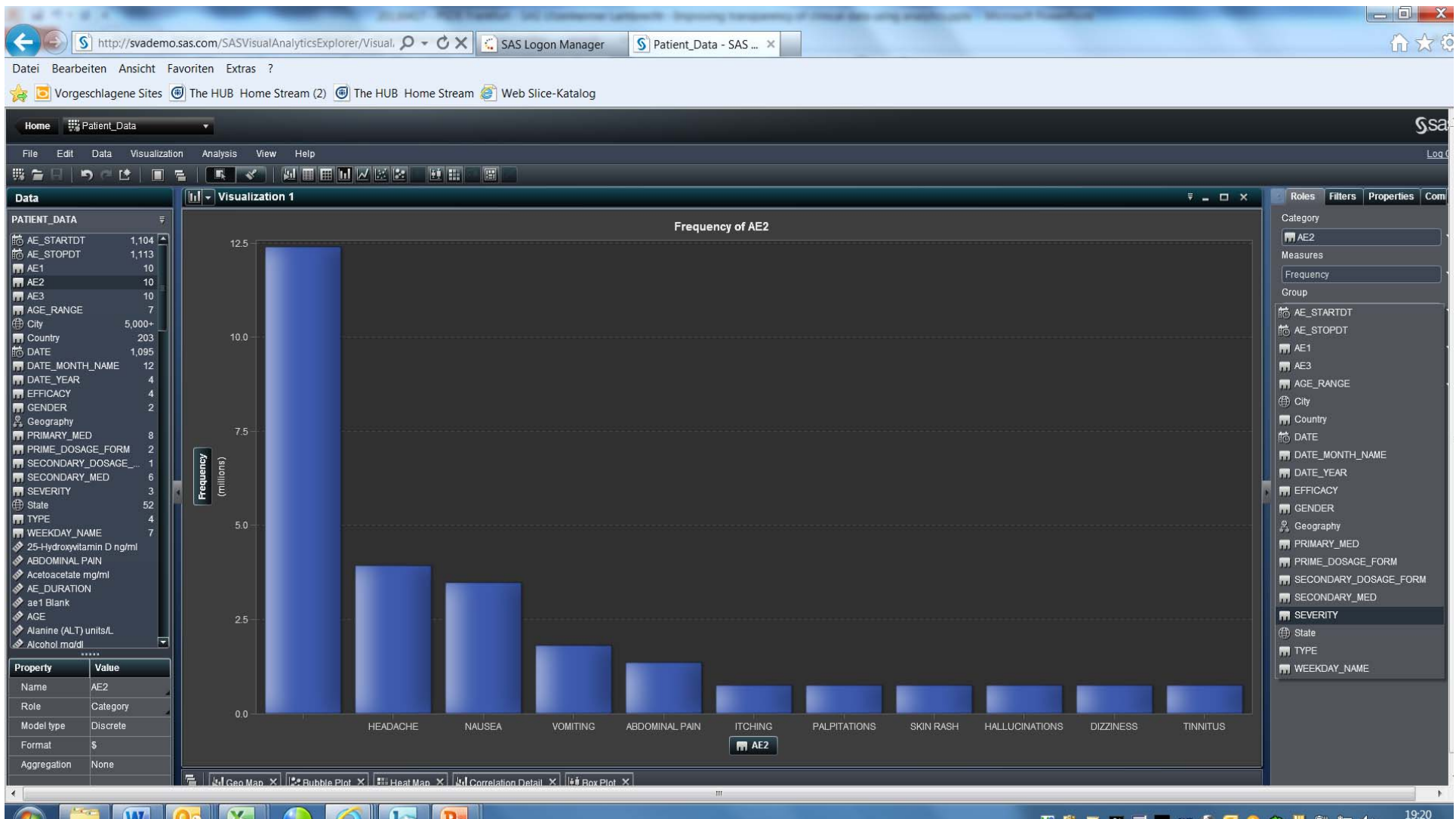
Right Panel: Filters Comments

Global Filters: Drop data items here to create global filters.

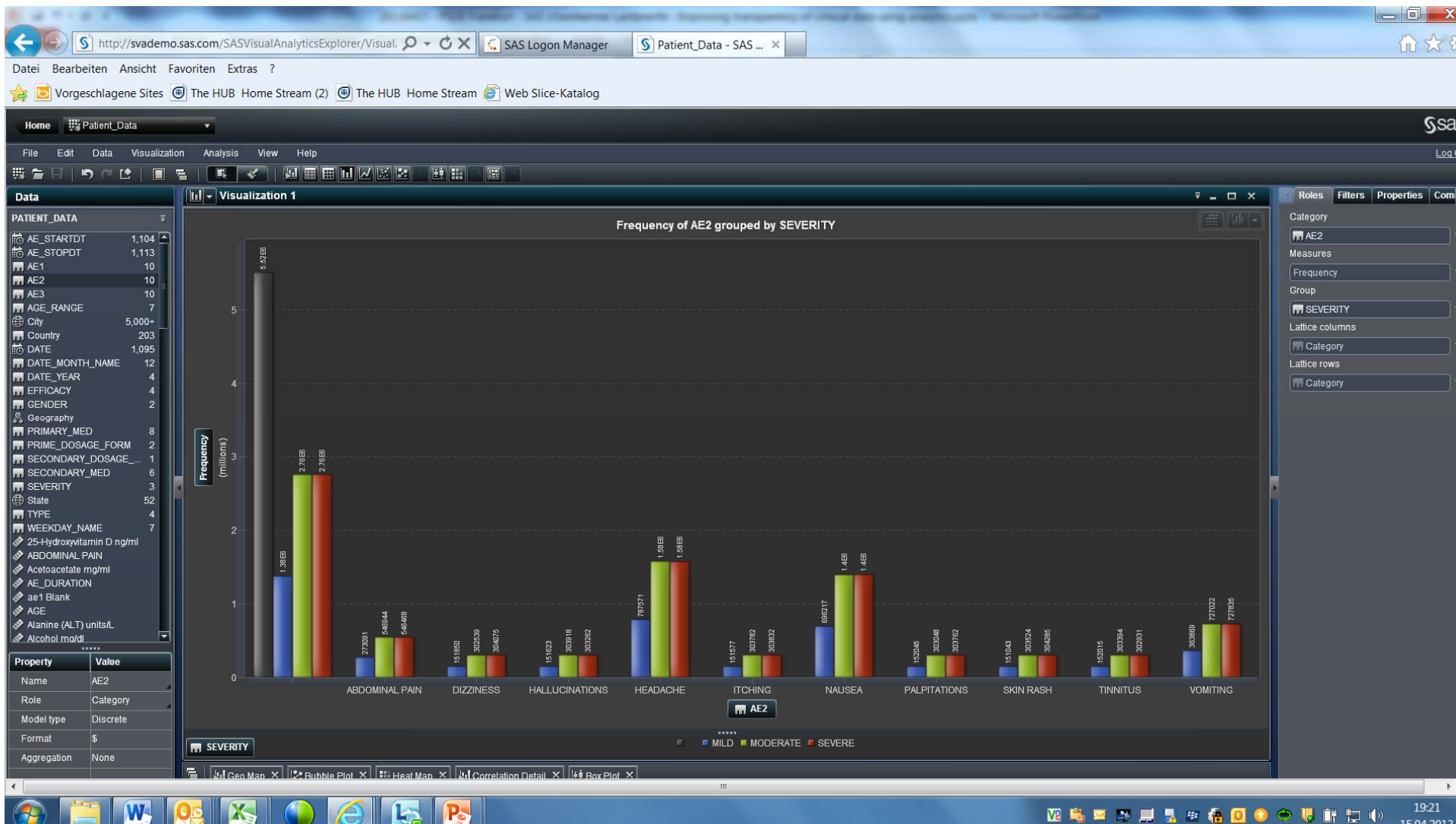
Bottom Panel: Geo Map, Bubble Plot, Heat Map, Correlation Detail, Box Plot

Taskbar: 19:19









Browser window showing the SAS Visual Analytics Explorer interface. The URL is <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual>. The interface includes a menu bar (Datei, Bearbeiten, Ansicht, Favoriten, Extras, ?), a toolbar with icons for file operations and analysis, and a main workspace area.

The left sidebar displays a list of data items under the heading "PATIENT_DATA". The items include:

- AE_STARTDT: 1,104
- AE_STOPDT: 1,113
- AE1: 10
- AE2: 10
- AE3: 10
- AGE_RANGE: 7
- City: 5,000+
- Country: 203
- DATE: 1,095
- DATE_MONTH_NAME: 12
- DATE_YEAR: 4
- EFFICACY: 4
- GENDER: 2
- Geography
- PRIMARY_MED: 8
- PRIME_DOSAGE_FORM: 2
- SECONDARY_DOSAGE_...: 1
- SECONDARY_MED: 6
- SEVERITY: 3
- State: 52
- TYPE: 4
- WEEKDAY_NAME: 7
- 25-Hydroxyvitamin D ng/ml
- ABDOMINAL PAIN
- Acetoacetate mg/ml
- AE_DURATION
- ae1 Blank
- AGE
- Alanine (ALT) units/L
- Alcohol mg/dl

The bottom section of the sidebar shows the "Property" and "Value" table for the selected item (AE2):

Property	Value
Name	AE2
Role	Category
Model type	Discrete
Format	\$
Aggregation	None

The main workspace area is currently empty, displaying the text "Drop a data item here". The right sidebar contains the "Filters" and "Comments" sections, with "Global Filters" indicating that data items can be dropped here to create global filters.

The bottom taskbar shows various application icons, including Internet Explorer, Word, Excel, and the Windows Start button. The system clock indicates the time is 19:23 on 15.04.2013.

Browser window showing the SAS Visual Analytics Explorer interface. The address bar displays <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual>. The page title is "Patient_Data - SAS ...".

The interface includes a menu bar (Datei, Bearbeiten, Ansicht, Favoriten, Extras, ?) and a toolbar. The main content area displays a list of data items under the "Data" tab, including:

- PATIENT_DATA
- AE_STARTDT (1,104)
- AE_STOPDT (1,113)
- AE1 (10)
- AE2 (10)
- AE3 (10)
- AGE_RANGE (7)
- City (5,000+)
- Country (203)
- DATE (1,095)
- DATE_MONTH_NAME (12)
- DATE_YEAR (4)
- EFFICACY (4)
- GENDER (2)
- Geography
- PRIMARY_MED (8)
- PRIME_DOSAGE_FORM (2)
- SECONDARY_DOSAGE_ (1)
- SECONDARY_MED (6)
- SEVERITY (3)
- State (52)
- TYPE (4)
- WEEKDAY_NAME (7)
- 25-Hydroxyvitamin D ng/ml
- ABDOMINAL PAIN
- Acetoacetate mg/ml
- AE_DURATION
- ae1 Blank
- AGE
- Alanine (ALT) units/L
- Alcohol mg/dl
- Alpha-glucosidase inhibitor
- Ammonia ug ofN/dl
- Amylase units/L
- Amylin Mimetic
- Ascorbic Acid mg/dl

A "Correlate Measures" dialog box is open, prompting the user to "Correlate these measures: (select at least 2)". The dialog lists the same measures as the main interface. The "Vertical axis:" and "Horizontal axis:" fields are empty. The "OK" and "Cancel" buttons are visible at the bottom of the dialog.

The bottom of the screen shows the Windows taskbar with various application icons and the system clock displaying 19:23 on 15.04.2013.

Browser tabs: <http://svademo.sas.com/SASVisualAnalyticsExplorer/Visual...> | SAS Logon Manager | Patient_Data - SAS ...

Navigation: Datei Bearbeiten Ansicht Favoriten Extras ? | Vorgeschlagene Sites | The HUB Home Stream (2) | The HUB Home Stream | Web Slice-Katalog

Application: Home Patient_Data

Menu: File Edit Data Visualization Analysis View Help

Left Panel (Data):

- PATIENT_DATA
- AE_STARTDT 1,104
- AE_STOPDT 1,113
- AE1 10
- AE2 10
- AE3 10
- AGE_RANGE 7
- City 5,000+
- Country 203
- DATE 1,095
- DATE_MONTH_NAME 12
- DATE_YEAR 4
- EFFICACY 4
- GENDER 2
- Geography
- PRIMARY_MED 8
- PRIME_DOSEAGE_FORM 2
- SECONDARY_DOSEAGE_... 1
- SECONDARY_MED 6
- SEVERITY 3
- State 52
- TYPE 4
- WEEKDAY_NAME 7
- 25-Hydroxyvitamin D ng/ml
- ABDOMINAL PAIN
- Acetoacetate mg/ml
- AE_DURATION
- ae1 Blank
- AGE
- Alanine (ALT) unit/L
- Alcohol mol/L

Bottom Left Table:

Property	Value
Name	AE2
Role	Category
Model type	Discrete
Format	\$
Aggregation	None

Correlate Measures Dialog:

Correlate these measures:
(select at least 2)

- ☐ Total Protein g/L
- ☐ Type_1
- ☒ Type_2
- ☐ Uric Acid mg/dL
- ☐ Urine_PH pH
- ☐ VISIT_NBR
- ☐ Vitamin A æg/dL
- ☐ VOMITING
- ☐ WEEKDAY_NBR
- ☒ WEIGHT
- ☐ White Blood Cell (WBC) k Cells/mm
- ☐ Zinc B-Zn æmol/L

Vertical axis: Glucose mg/dl, Hemoglobin A1c %, Type_2, WEIGHT, BMI

Horizontal axis: Bilirubin (Total) mg/dl, Glucose mg/dl, Hemoglobin A1c %, Type_2, WEIGHT

Buttons: OK Cancel

Right Panel: Filters | Comments

Global Filters

Drop data items here to create global filters.

Taskbar: 19:25 15.04.2013

