



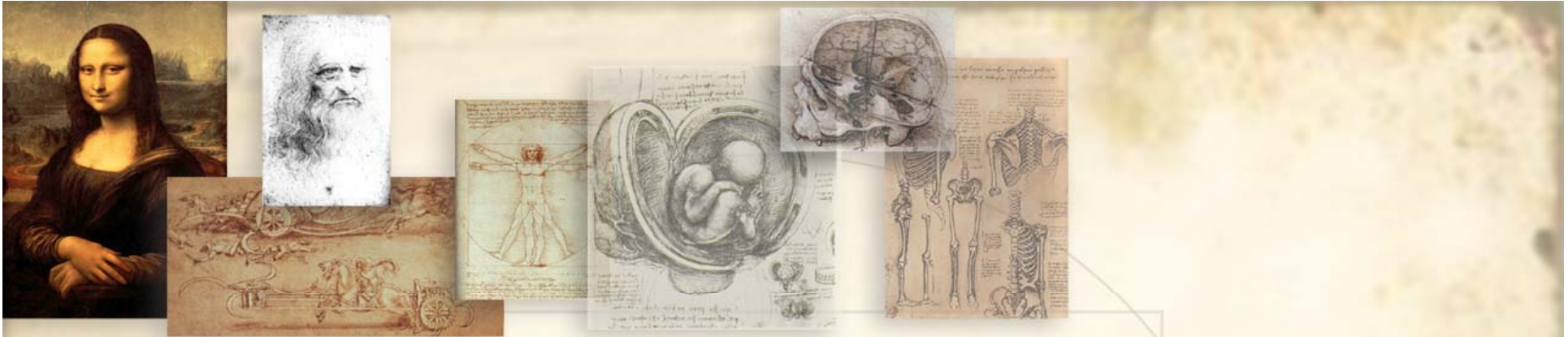
# Applied Market Research and Analysis

Introductions

Basic Problems

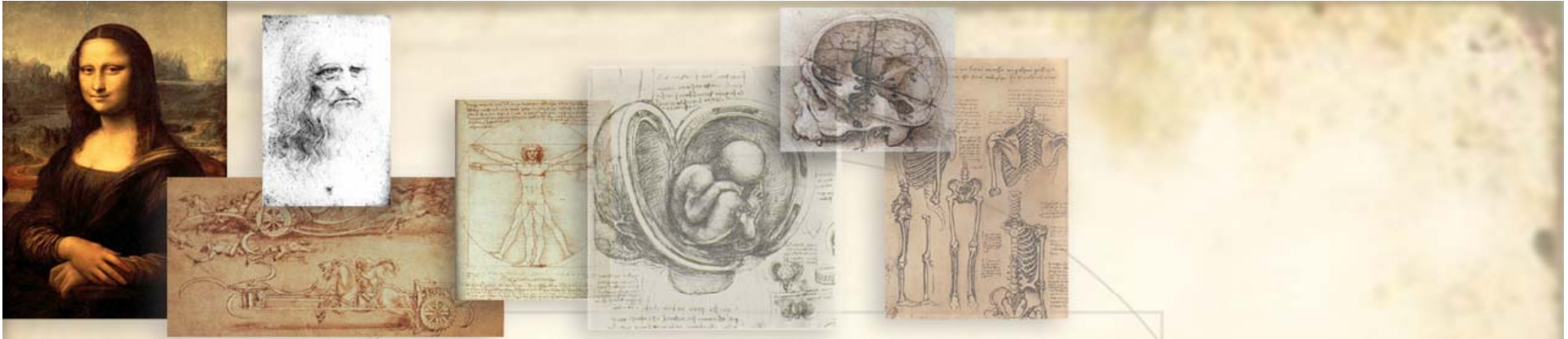
Research Process

Inductive/Deductive Research



# James Reardon, PhD (Jim)

- Academic
  - BS (CIS), MBA, PhD (Mktg)
  - UofTN, UNC, UofAntwerp, KIMEP, UofLju
  - Research/Scholarship
- Industry
  - Own/Mgmt
    - Construction, Retail, Casino, Film Distribution, E-com
  - Mgmt – TV shopping, Consulting
- Personal



# Introductions

- Name
- Industry experience
- Goals – i.e., why are you here?
- Interesting



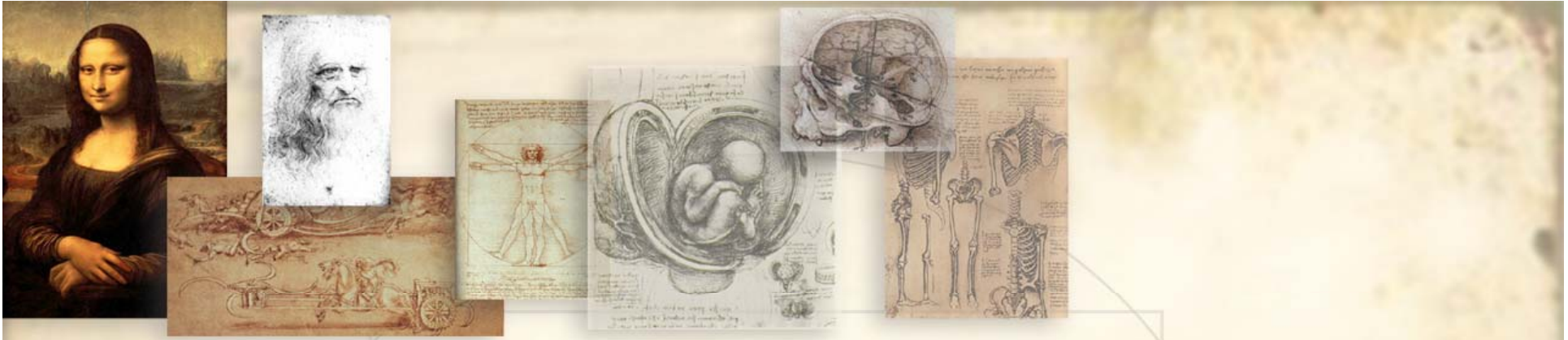
# Problems Requiring Research

e.g. A.C.N.

Putting Out Fires

VS

Solving Underlying Problems



# Why do research?

Topic	What we typically know	What we usually don't
Opportunity Analysis (GAP/SWOT)		
Product		
Price		
Distribution		
Communication		
Raw Materials		
Producer		
Wholesaler		
Retailer		



# Research Process

1

- Determine Research Problem

2

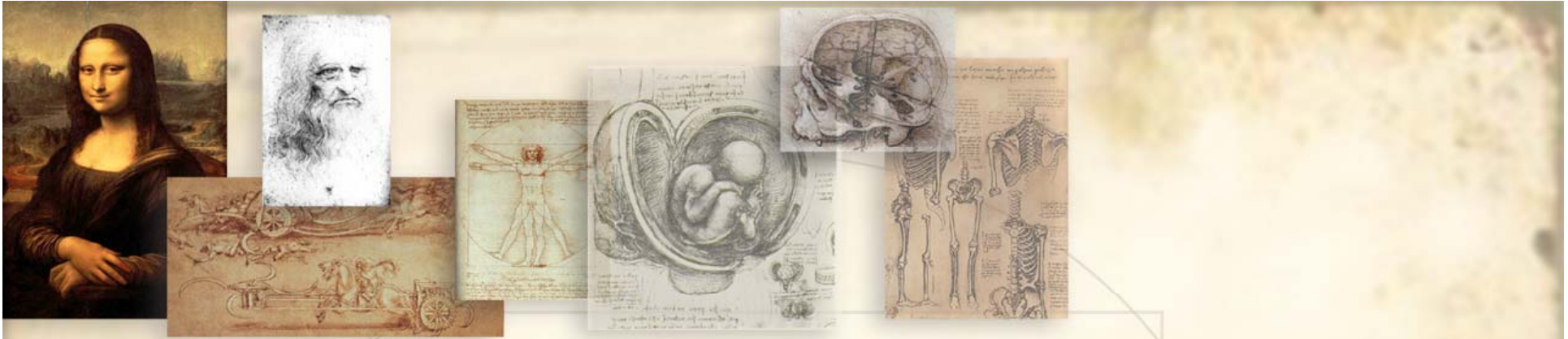
- Select Appropriate Design

3

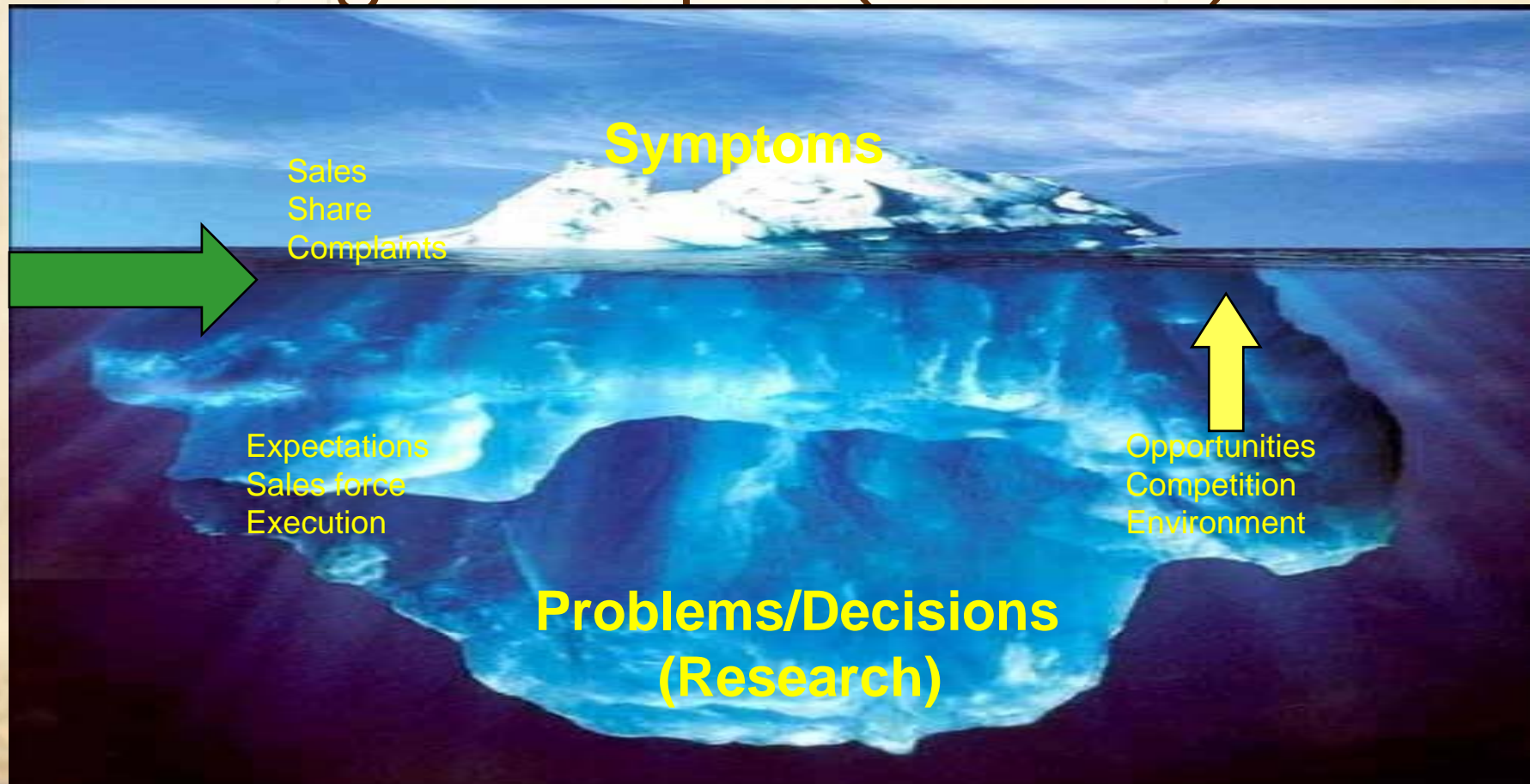
- Execute Research

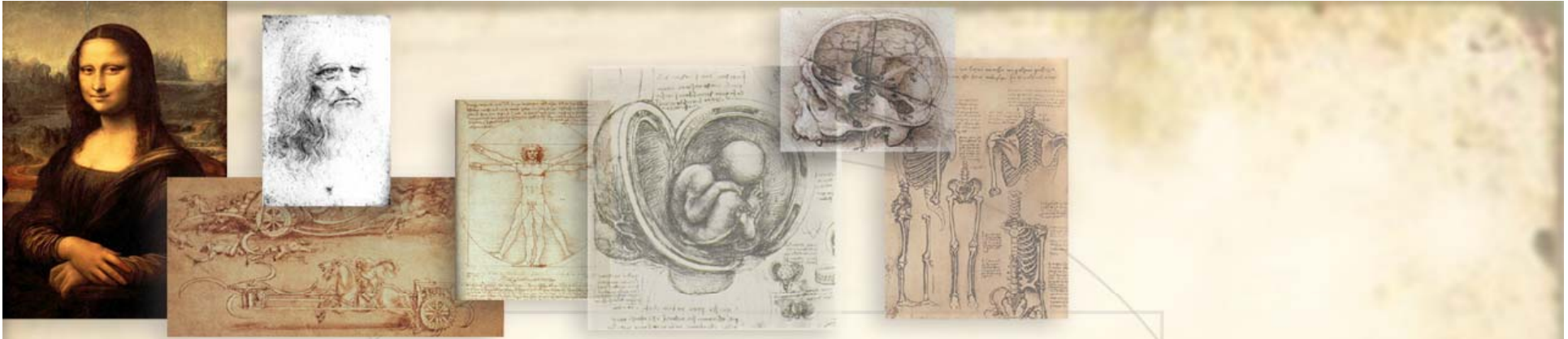
4

- Communicate Results



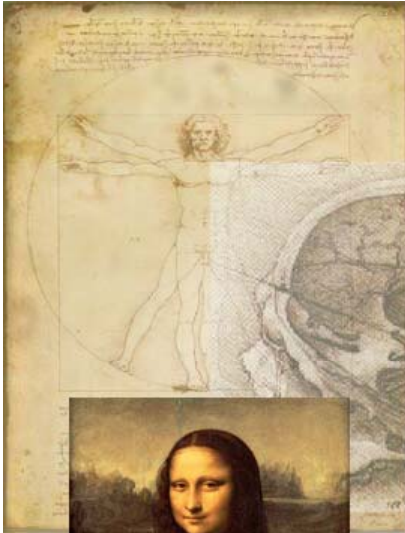
# Iceberg Principle (Exh 2.5)





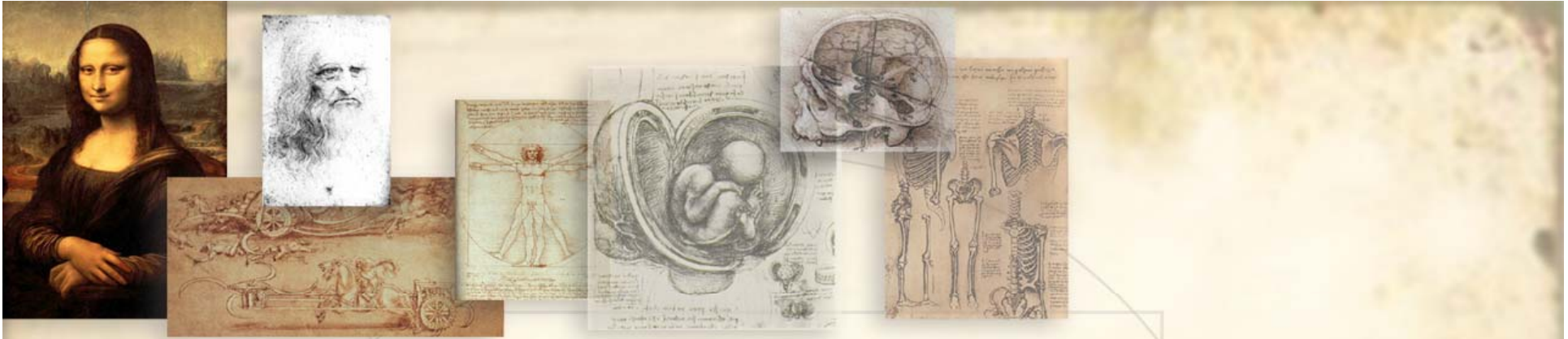
## Assignment for Today

- What is your research question?
- Is it a managerial or research problem? Identify both for the problem at hand.

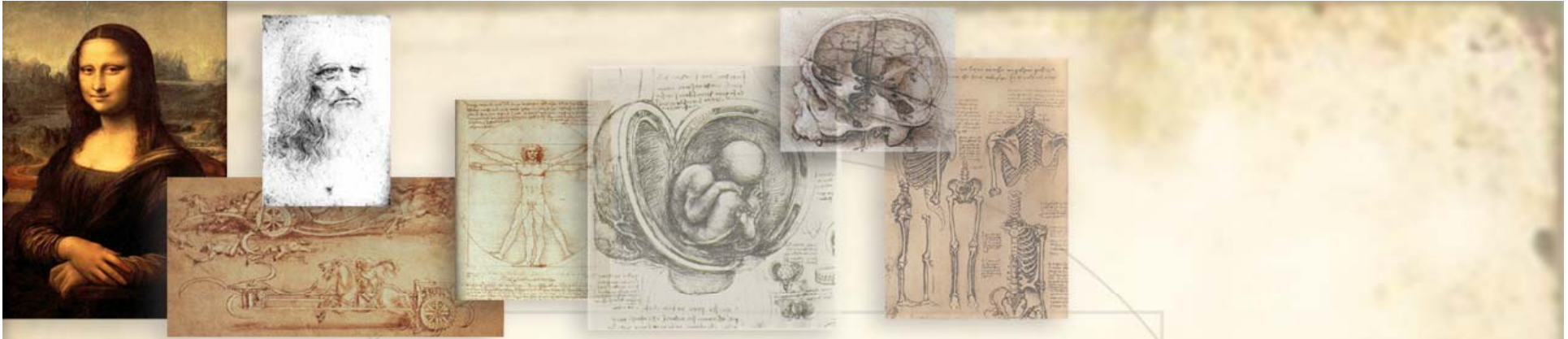


# Deductive and Inductive Research

Causal  
Predictive  
Descriptive

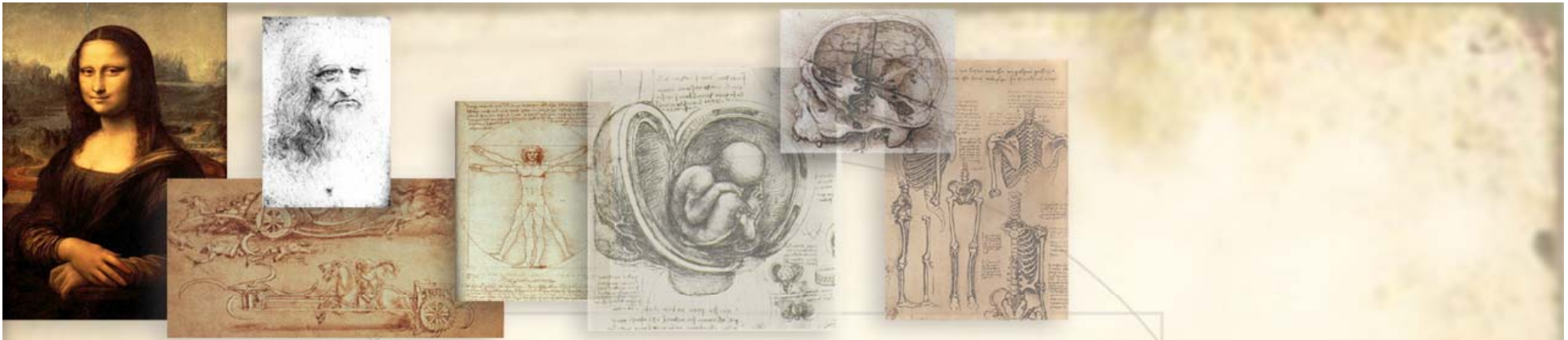


- **Define**
  - Causality – Time, Correlation, and Logic
  - Prediction – Correlation and Time?
  - Description – Situation as it exists
- **Processes**
  - Causality – Review, Hypothesis, Test
  - Prediction – Test
  - Description – Show



## So which do we use?

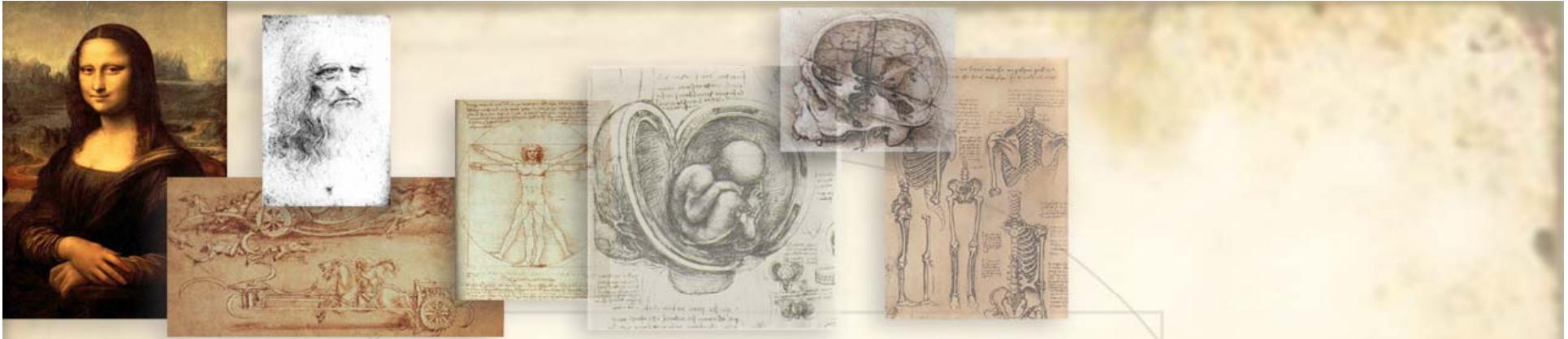
- Purpose – make an informed decision at limited risk/exposure
- Cost versus Benefits
  - Benefits
    - Revenue, Market Share, Strategic
  - Costs
    - Time – Labor/Talent and Opportunity cost
    - Budget
    - Risk – importance
    - Image/Brand Equity



# Which do I use? (loose criteria)

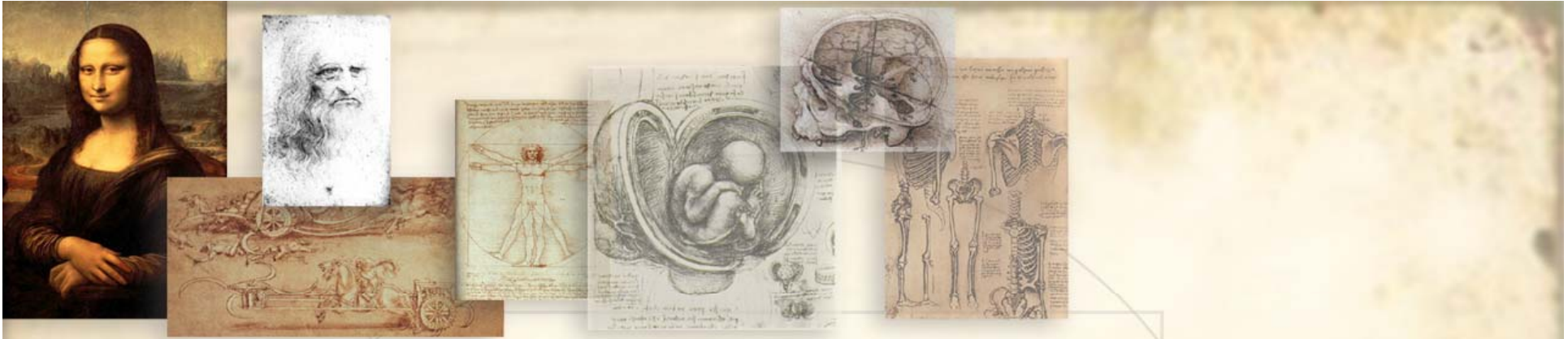
Buying Television Time (\$50Mil+ per year 200+ contracts)

Criteria	Descriptive Info for Contract	Predictive Modeling Est. Sales	Causal Reexamine Model
Cost	< \$100K	> \$200K	\$5M+
Relationship w/Seller	Excellent/Good	Poor	None
Market Size/Charac	Similar to Existing	New	Very Large/Unusual
Type (Broadcast, Cable, Satellite)	Broadcast	Any	New Type
Channel #	99, 1-23	24-60	100+
Other	Company Strategy		Unusual



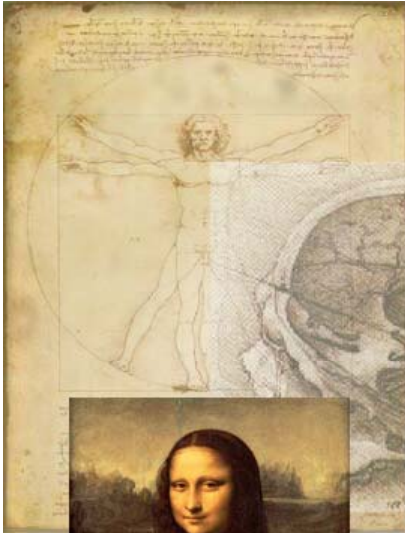
## When not to use research

- Time/Budget restrictions
  - Limited time opportunities – closeouts
- The answer is obvious (be careful)
- Limited value of decision
- Information is trendy but not used
- Give the competitor information
  - Product Launches



## Other uses of Mktg. Research

- Idea Generation
- Internal – Mgmt. etc..
- Questionable Uses
  - Self Promotion – Ethical? Research?
  - “Proof of Concept” i.e., Political

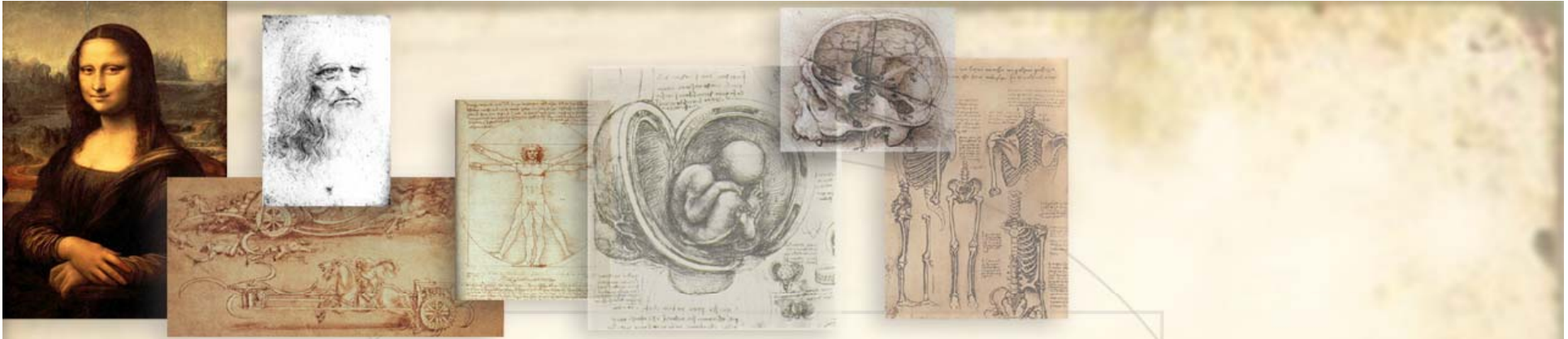


## Secondary Data (being efficient)

Definitions

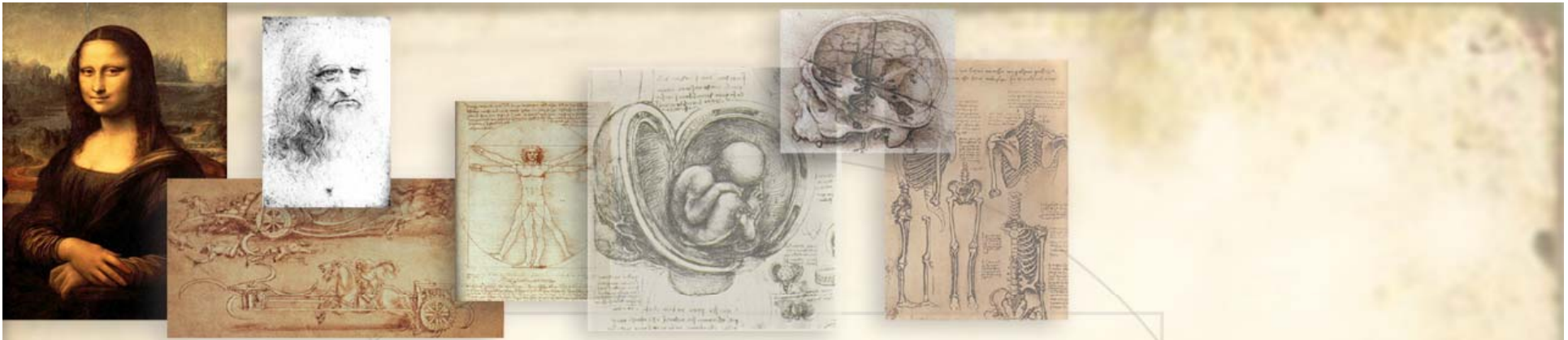
Internal vs.

External Sources



## Definitions

- Secondary – Gathered not for the purposes at hand
  - Internal – Gathered by company
    - Note: Some data w/i company is not gathered by the company – e.g., compiled databases
  - External – Gathered by someone else



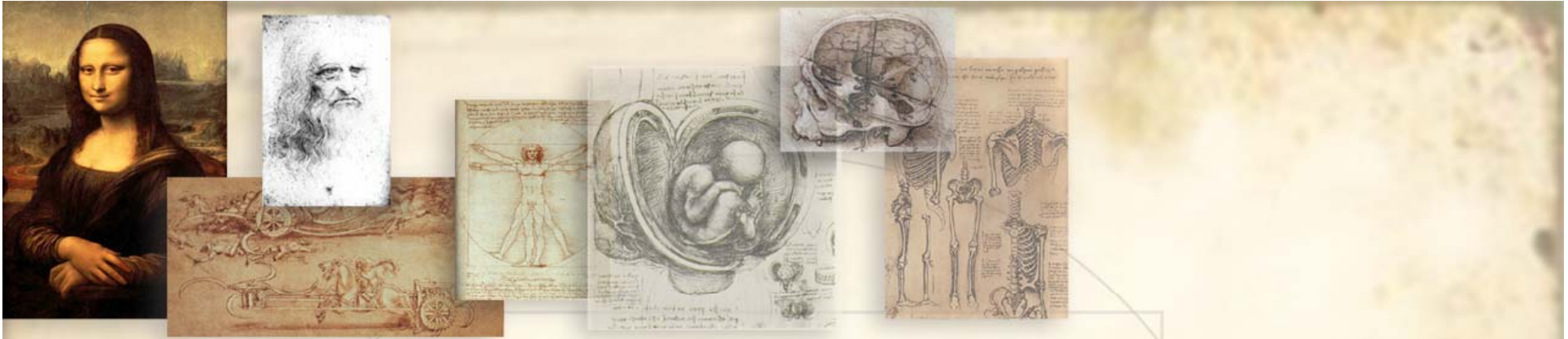
# Why/Why Not Secondary Info

- **Benefits**

- Speed
- Cost
- Access
- Ease

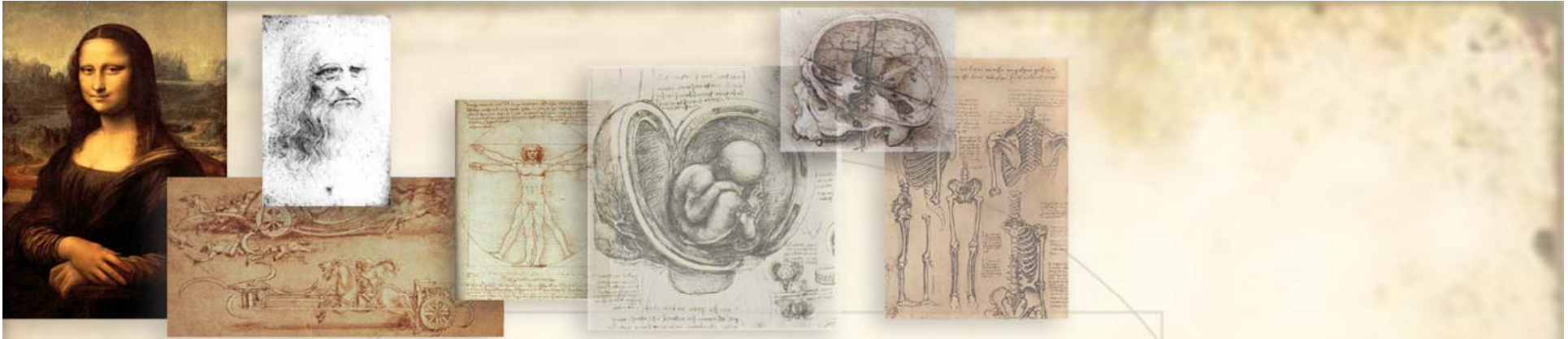
- **Costs**

- Fit/Purpose
- Accuracy
- Consistency
- Credibility
- Methodology
- Bias



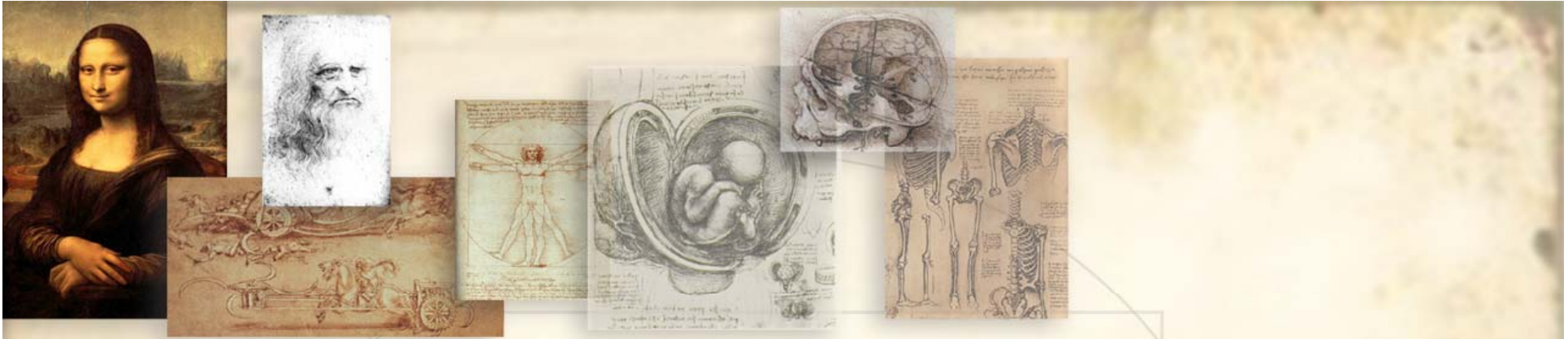
## Internal Data (Exhibit 4.1-4.2, p.64-5)

- Cheapest (Time and Money)
- Most Accessible
  - Who are your friends at work?
- Accuracy - Best knowledge of source/procedures for compilation



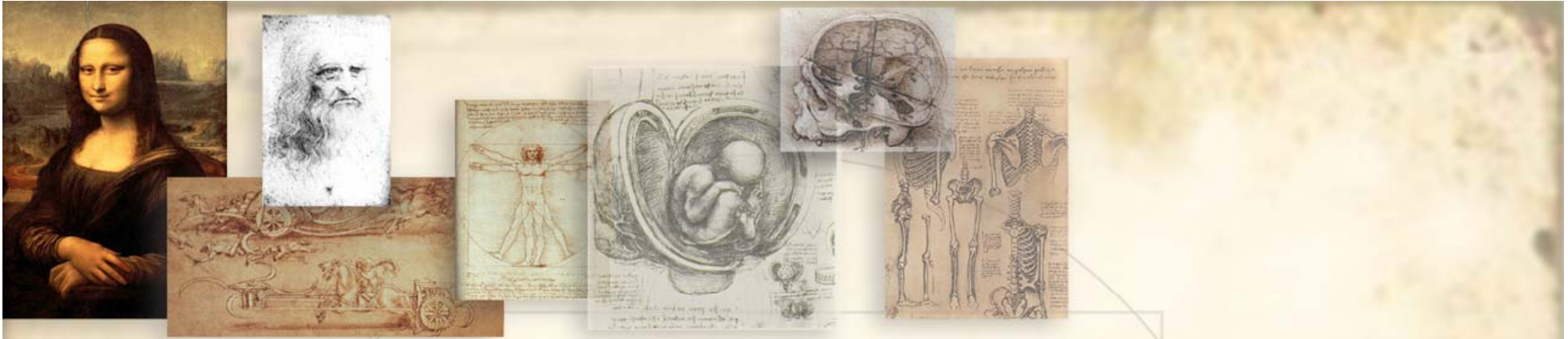
## Example of Secondary data (External)

- [http://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_GDP\\_\(nominal\)](http://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal))
  - Italy ½% off – also see #12/13 rank
- [http://www.indexmundi.com/italy/gdp\\_real\\_growth\\_rate.html](http://www.indexmundi.com/italy/gdp_real_growth_rate.html) -- 0.10%
- <http://www.imf.org/external/np/ms/2006/111306.htm> -- 1.75%
- <https://www.cia.gov/library/publications/the-world-factbook/geos/it.html> -- 1.9%
- **Why?**

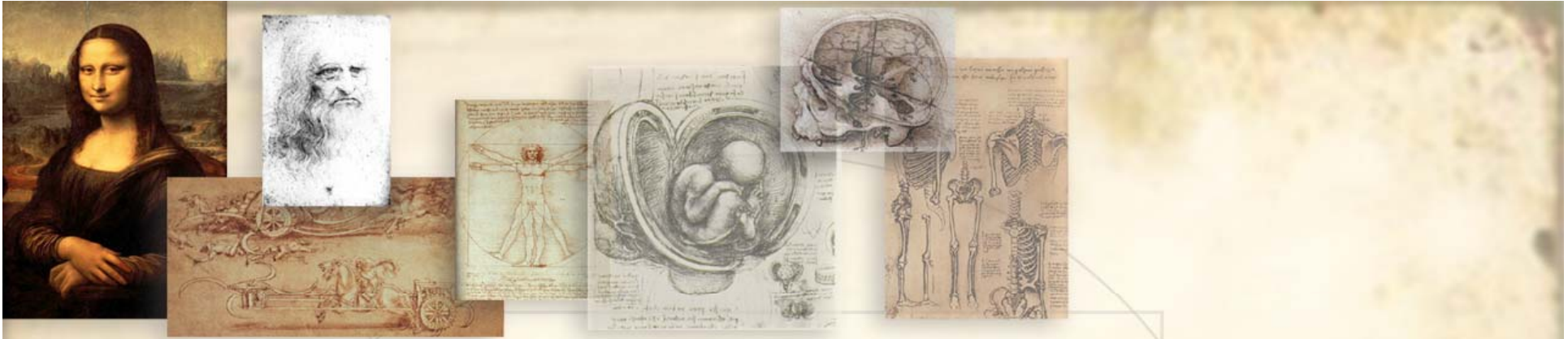


## External Data (by reliability)

- **Commercial/Syndicated**
  - Live and Die by accuracy of information
  - No inherent bias to produce
  - <http://www.knowthis.com/market-research/leading-marketing-research-companies.htm>
  - AC Nielsen [http://acnielsen.com/company/about\\_nielsen.shtml](http://acnielsen.com/company/about_nielsen.shtml)
  - Euromonitor <http://www.euromonitor.com/>
  - GIS systems – combine proprietary software with government data



- **Government/Related Orgs**
  - Limited rationale for bias of data
  - Limited incentive for accuracy
    - UN and Related
    - EU
    - OECD
    - Governments
      - Usually Ministry of Economics/Finance
      - US – Department of Commerce

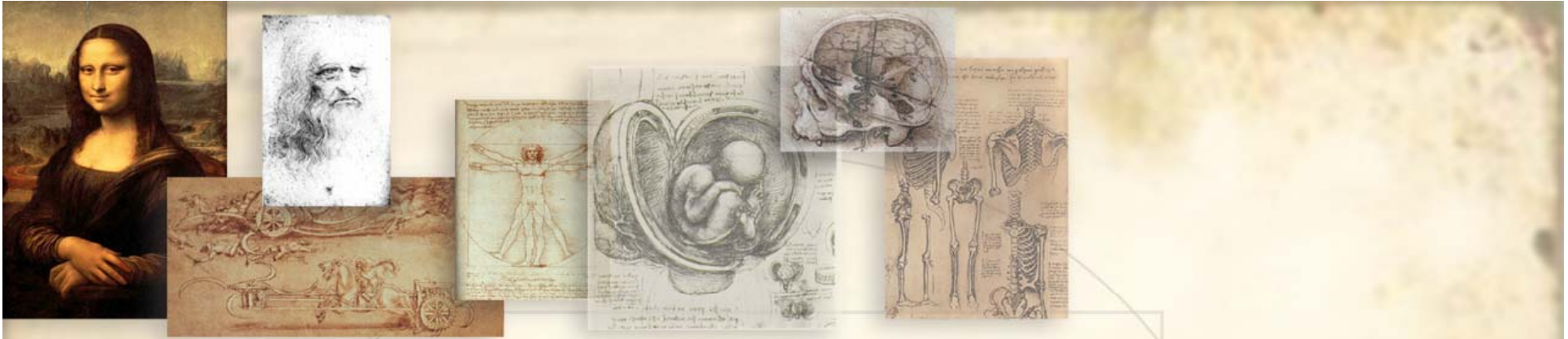


- Trade/Industry Groups
  - Some rationale for bias of data
    - PR for Industry (e.g., Tobacco)
  - Limited incentive for accuracy
    - Dependent on members for volunteer information
  - For almost all industries/trades/hobbies
    - <http://www.marketingsource.com/associations/information.html>
    - <http://gale.com/servlet/BrowseSeriesServlet?region=9&imprint=000&cf=ps&titleCode=EA1&edition>





# Tuesday Assignment

- Read as per course outline
- Refine your research question.
- What type of research?
  - See previous table
- Suggest potential causes.
  - Use simple logic for now



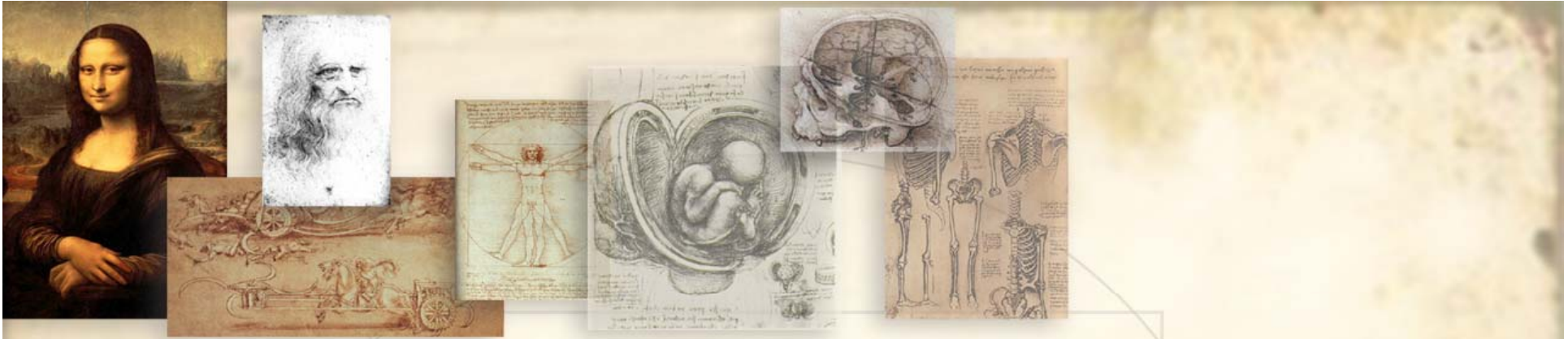
## Assignment (continued)

- Identify 2-4 Exploratory sources to further define or answer your R Q
- Find information related to your question as available.



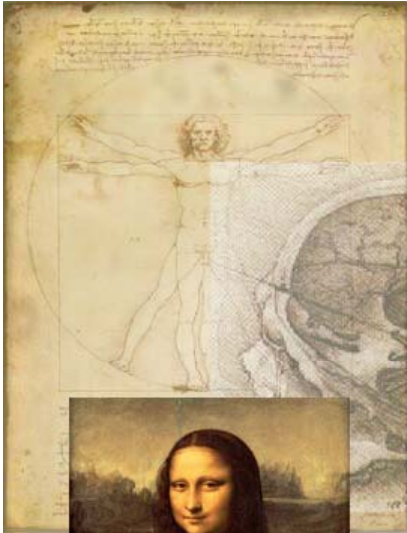
# Applied Marketing Research and Analysis Tuesday 18-3-08

Finding Evidence  
to Solve Problems



# Assignments

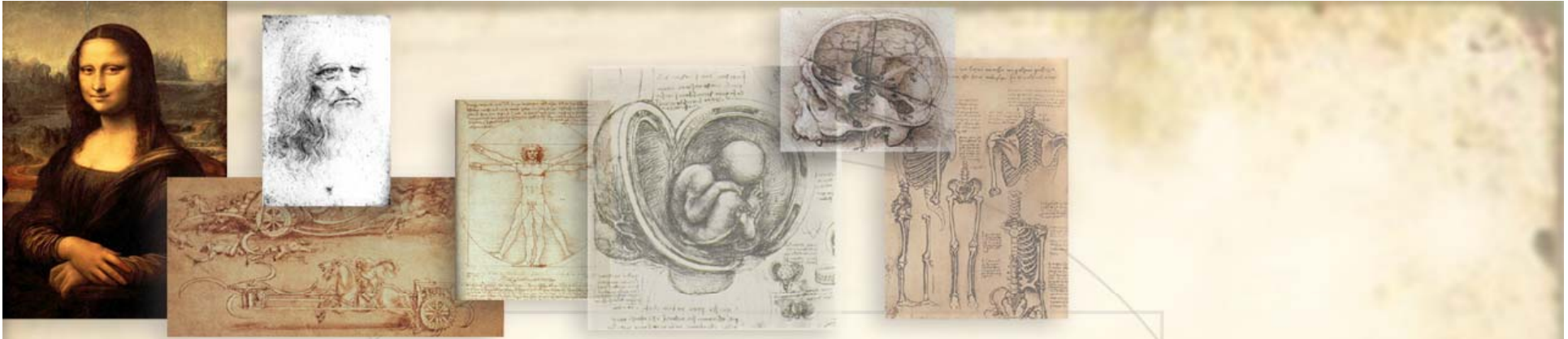
- Reformulated Question
- Descriptive, Predictive or Causal?
  - Why?
- Potential Answers/Causes?
- Secondary sources



# Qualitative/Exploratory Overview

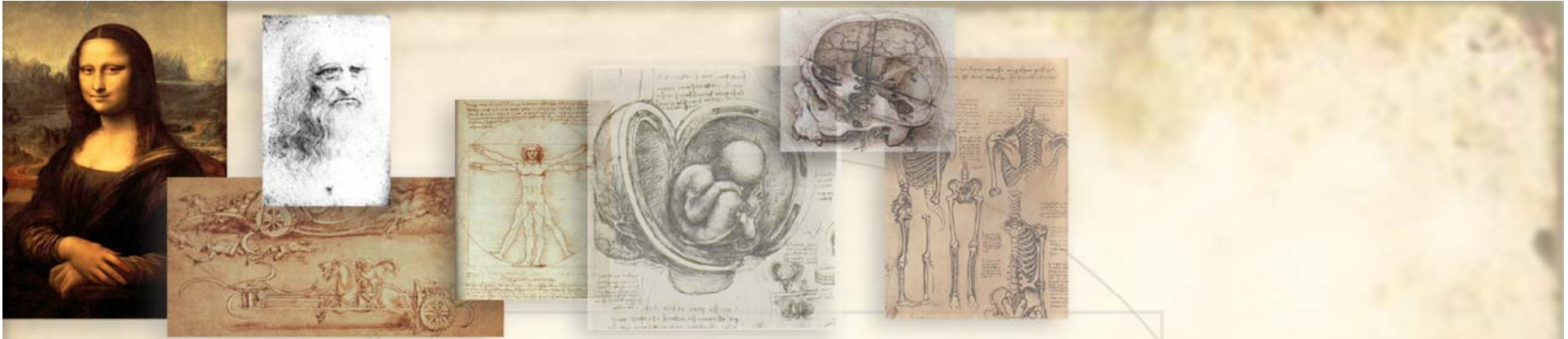
Purpose

Types/Procedures



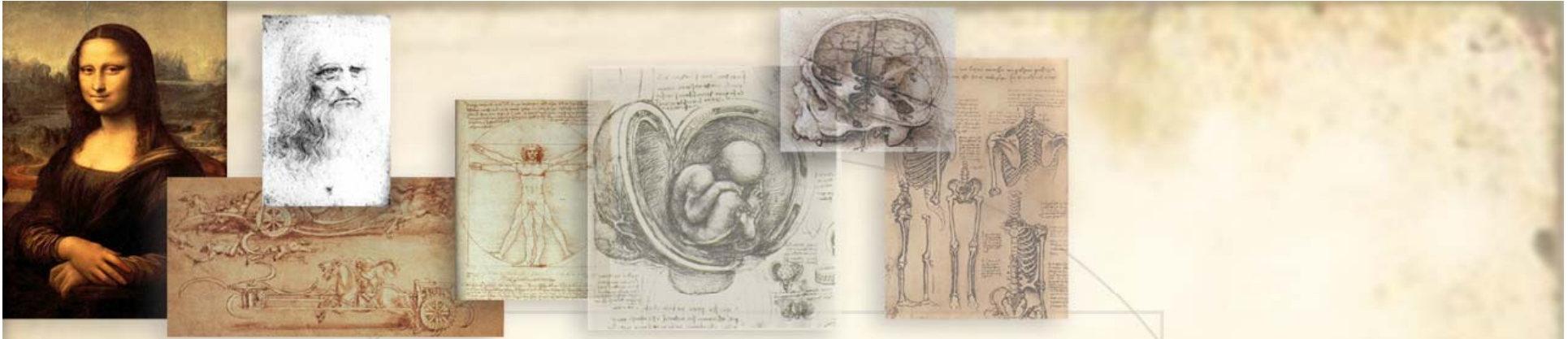
e.g., Clothing in Argentina

- **Clothing Designer/Producer**
  - Very successful
  - Wanted to export to France



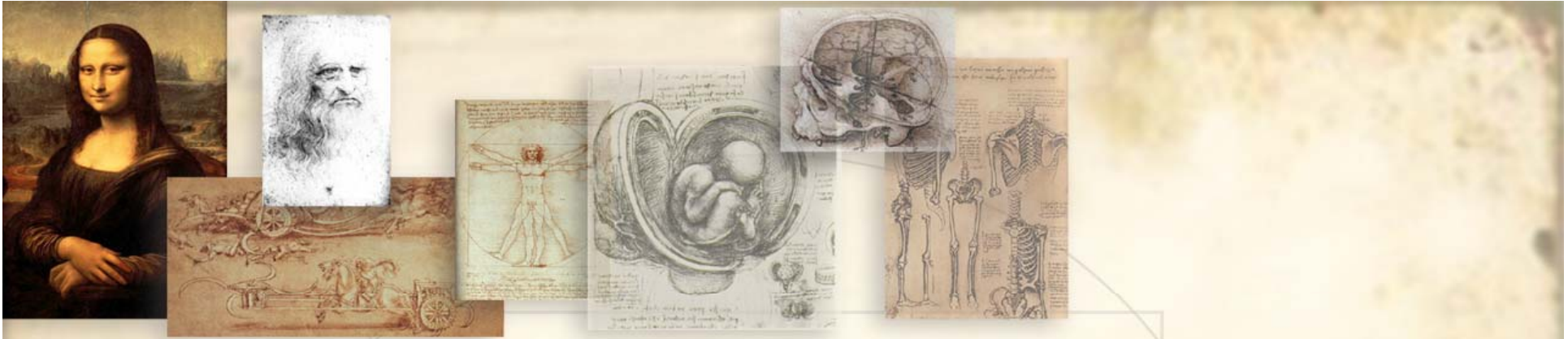
# Literature Review

- Should always start here after internal data – easy, quick and inexpensive
  - Academic Literature – personal view
  - Trade Rags
  - Case Studies – ideas
  - General Internet



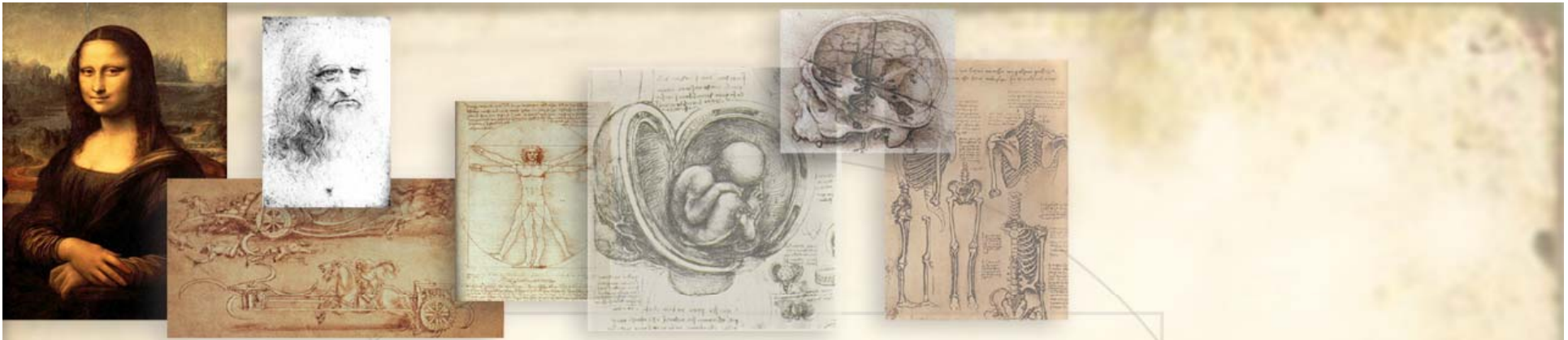
## Types of Qualitative/ Exploratory

- **Key Informant/Experience/Depth Surveys/Interviews**
  - Want to know the answer.....
  - Sometimes the shortest path is most obvious/direct
  - KI/Experience differ slightly from Depth



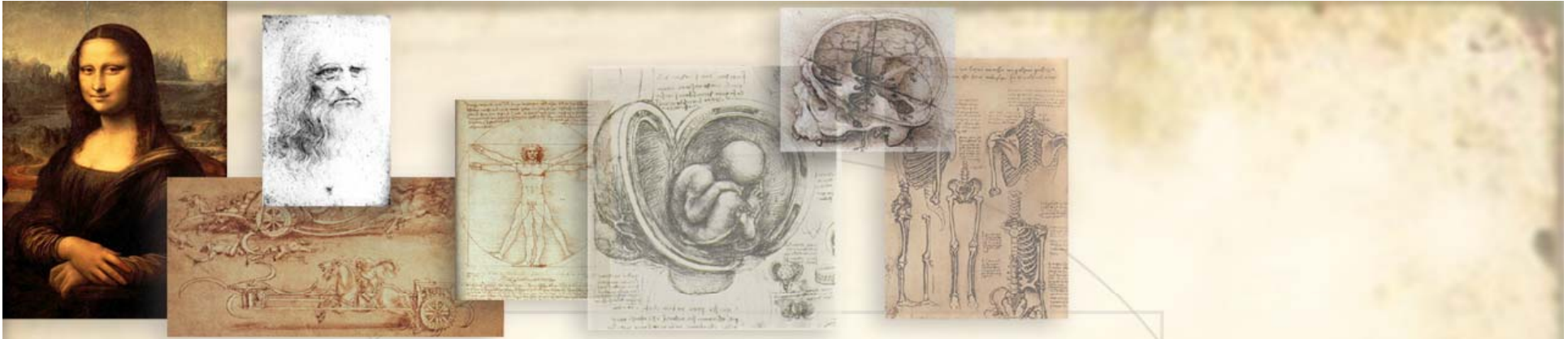
## Focus Groups (Exhibit 5.5)

- Most popular method
- Moderator is key
  - Not control, but guide and record
- 8-12 People – homogenous group
- 4-8 different groups - heterogeneity
- 1.5-2 hours



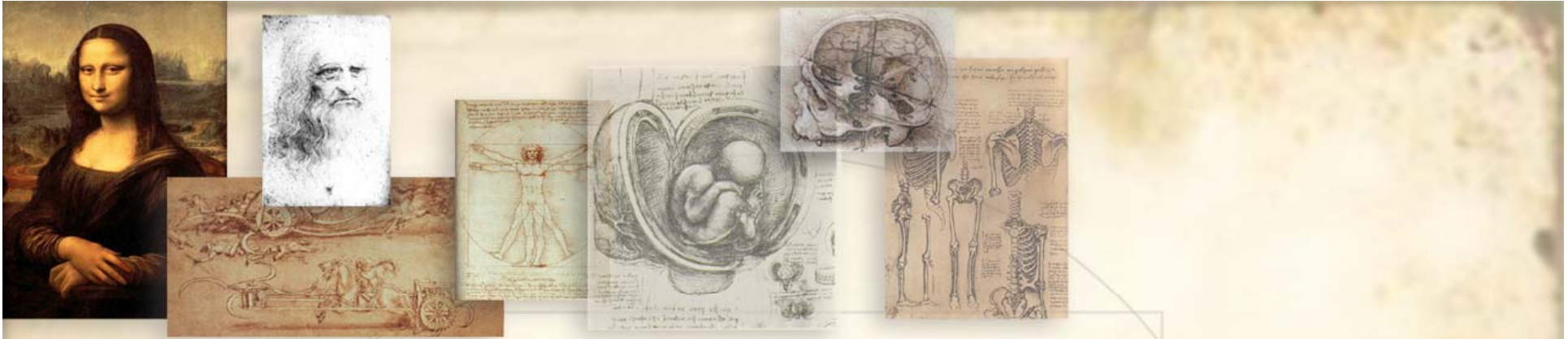
# Observation

- **Ethnography**
  - Watch and see
  - Some activities are highly affected by Hawthorne effect
- **Netnography**
  - Artificial Environment of net –
  - Alt. or discussion groups



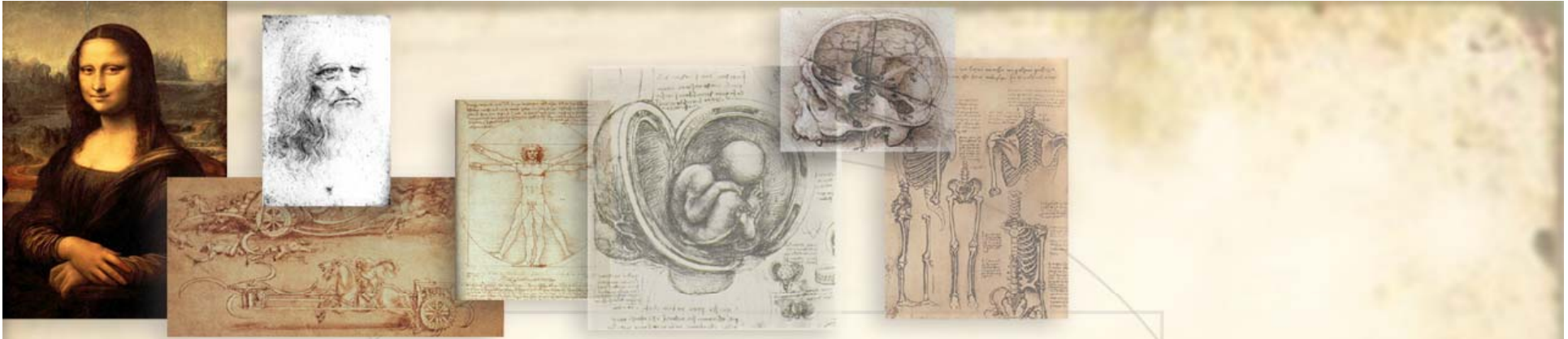
# Projective Techniques

- Word Associations
- Sentence Completion
- Group story telling
- TAT – tell me a story from a picture



# New Projective

- ZMET –
  - Individually Gather Pics of subject
  - As a Group
    - Compare/contrast
    - Create a story/movie
    - Summarize them into an image



# Assignment for Wed 19-3-08

## Part I

- Identify 2-4 Qualitative techniques to further define or answer your R Q
- Develop the materials necessary to conduct
  - Who?
  - Guides
  - Data Capture/Forms



# Primary Research Quantitative Overview

Measurement Basics

Survey Design

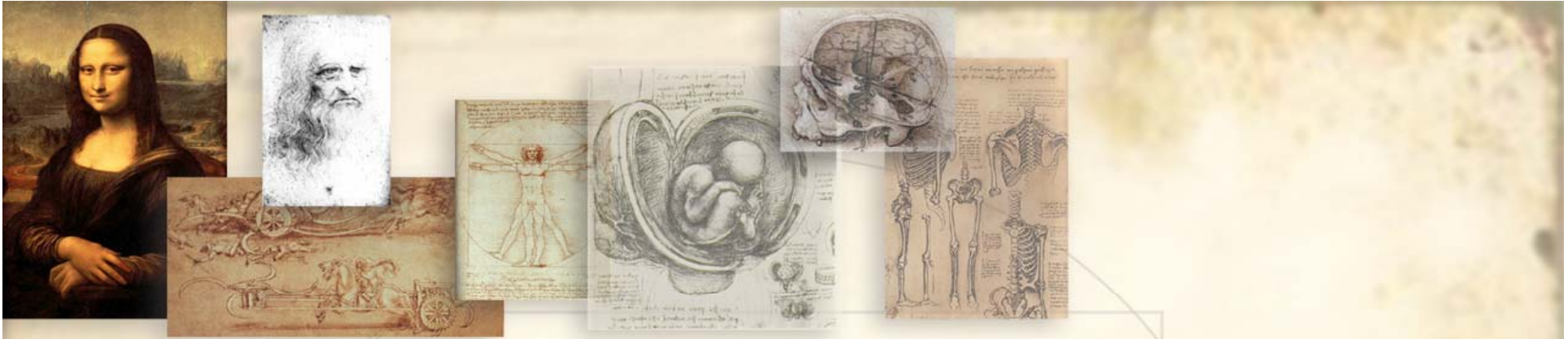


# Measurement Basics

Definitions

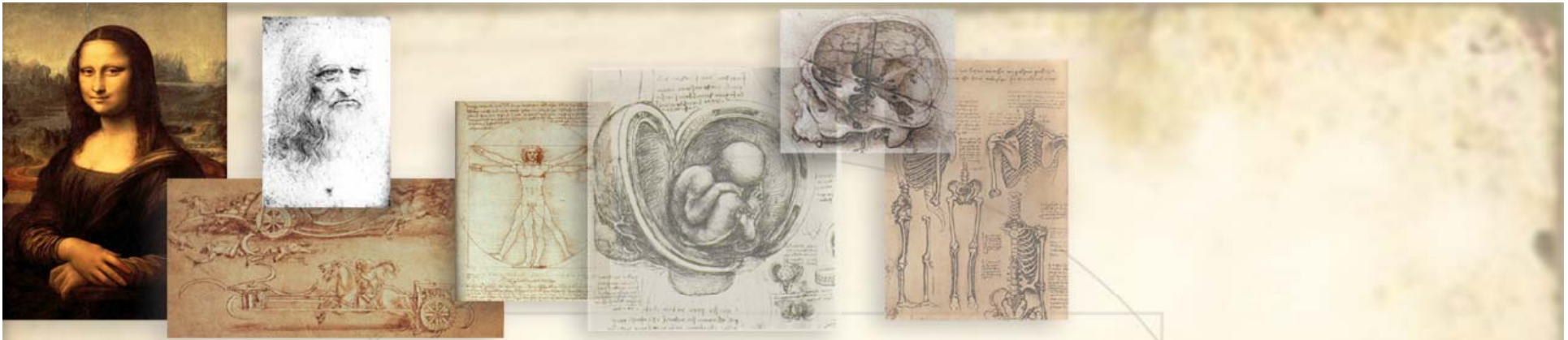
Rules of Thumb





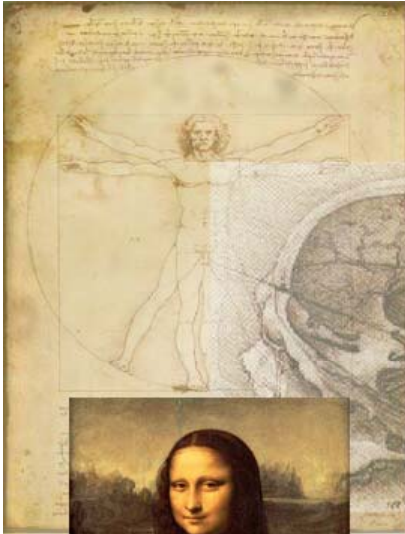
# Measurement Basics

- Important when developing questions
- Type of Data
  - Nominal, Ordinal, Interval, Ratio
  - Binary – some special rules apply
- ROTs
  - Use highest level data possible
  - Don't over-specify
    - meters w/ micrometers



# Measurement Levels

Scale	Types	Assigning Rules	Application	Stats
Nominal	Binary, Category	Objects either identical or different	Classification	Mode, %, Chi-Squared
Ordinal/Rank	Comparative, Rank, Paired Comparison	Objects greater or smaller than other	Rankings, relative preferences,	Percentile, Median, S. Correlation
Interval	Likert, Stapel, SD	Equal intervals no absolute 0	Attitude scales temperature	Mean, nearly all stats
Ratio	Physical, money	Same as above except 0 has meaning	Sales, income, production etc.	Mean (geometric), all stats



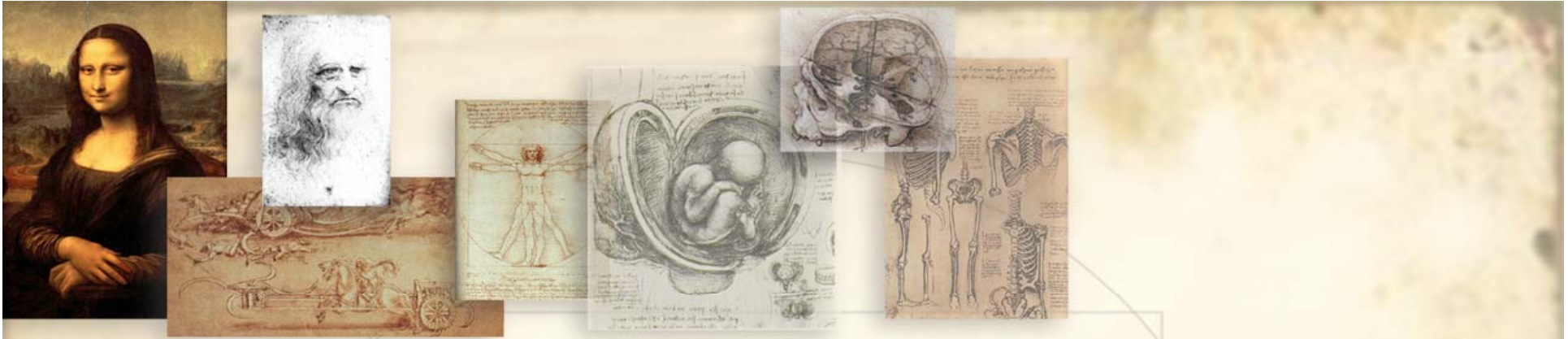
# Survey Design

## How (not) to ask a Q

Types

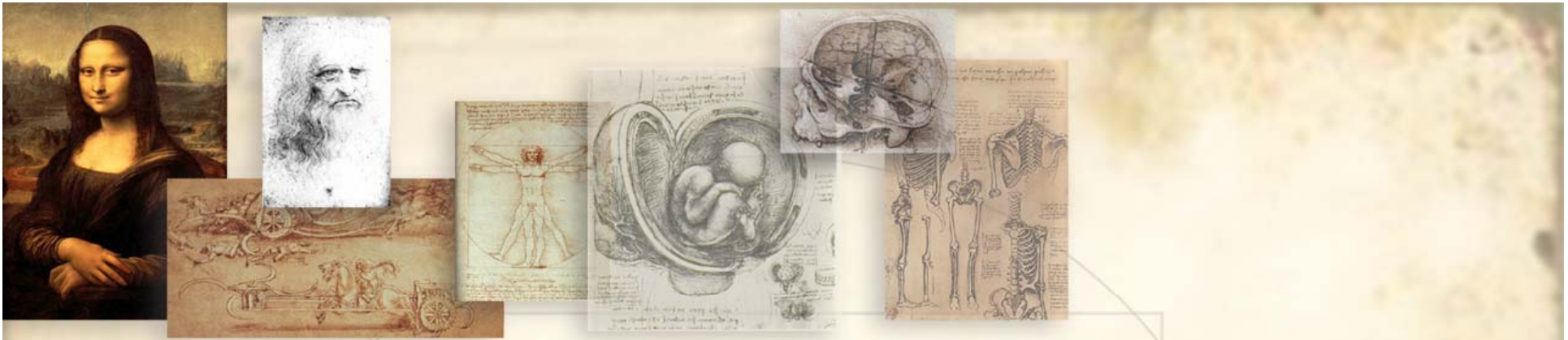
Format

Presentation



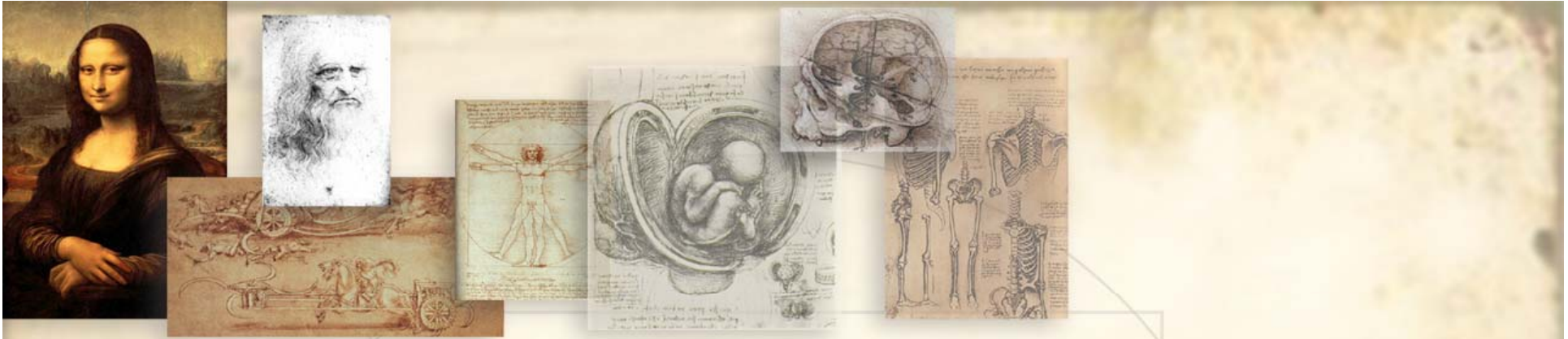
When in doubt – refer back to research question/problem

- Create “data needs” and rank them
  - Absolutely need
  - Probably need
  - Would be nice
  - Interesting
- If this is done poorly, everything will be



# Data Needs

Research Q	Variable Name	Use	Units	Unit/Type
Outcomes		Measure effect		
Sales	GSales	Outcome	000 Euro	Ratio
Hypothesized Antecedents		Check antecedents		
Ad Expend	AdExpen	Effect	000 Euro and lagged -1 to -4	Ratio
Predictive		Related, not causal		
Season	Fall, Spring, Winter, Summer	Predictive	0 if not season, 1 if season	Binary

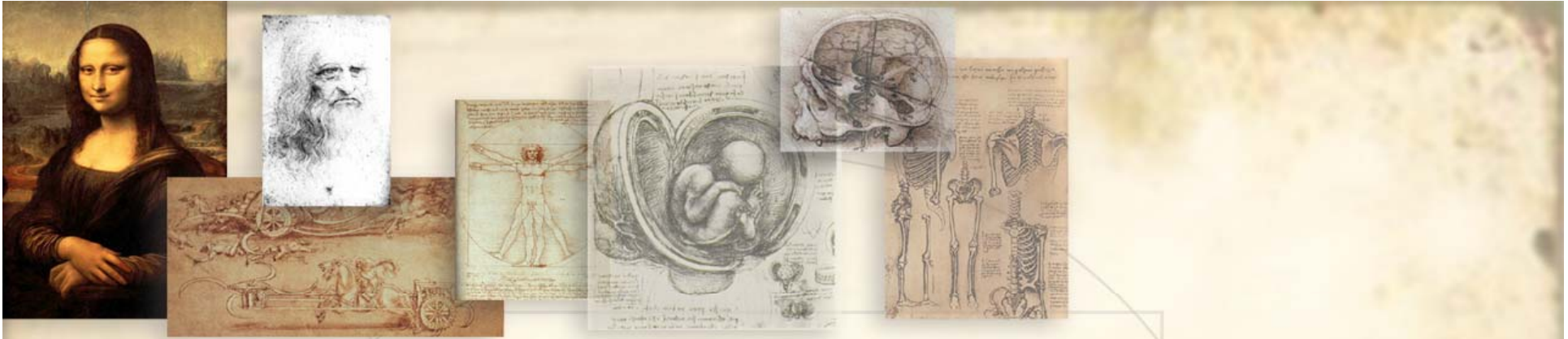


## Types of Survey Collection Exhibit 6.2

- Personal Interviews
  - Home vs. Mall Intercept
- Telephone
- Self Administered
  - Mail, Online, Drop-offs

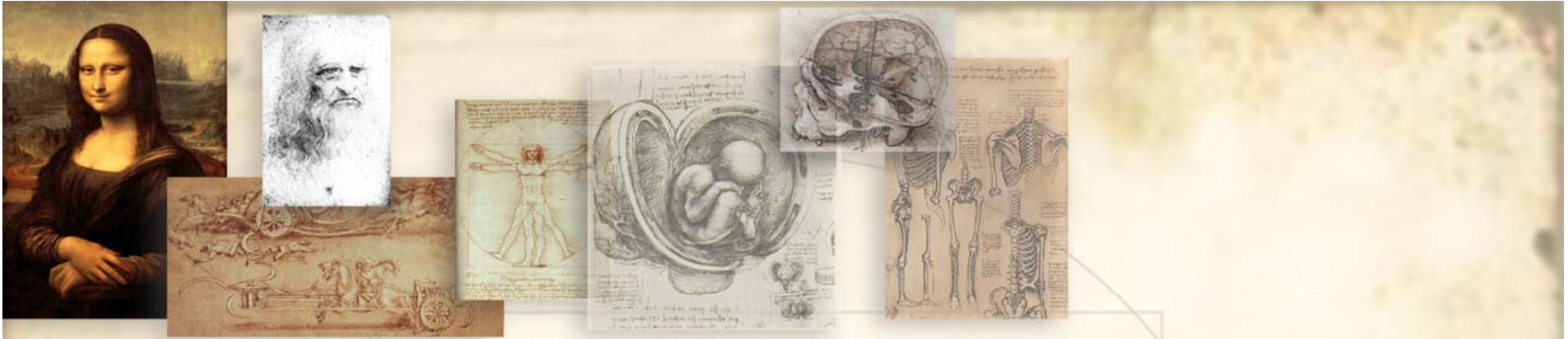


	<b>Personal</b>	<b>Telephone</b>	<b>Self Admin</b>
Cost	Highest	High	Low
Speed	Fast	Fastest	Slow
Adaptability (stimuli, flow)	Best	Good	Limited
Anonymity	Limited	None	Best
Monitoring Ability	OK	Best	Limited
Data Quality	Best	Good	Limited
Complications	Interviewer bias	Non-Response Legal No home phone	Non-Response
Used best for:	Depth	Politics	Personal/ Confidential



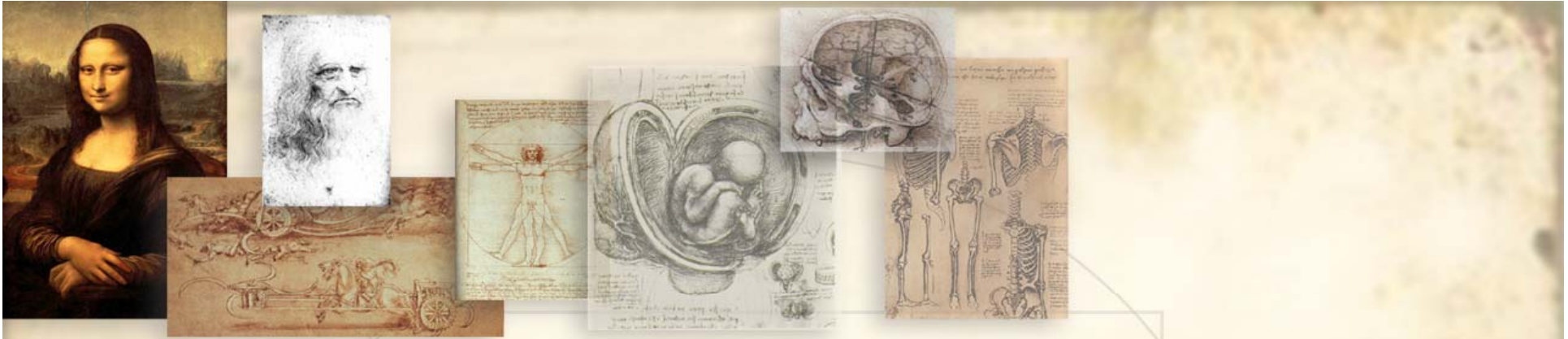
## Errors to Avoid

- Interviewer Bias
  - Demographics that create the highest response rate will also create the highest bias.
- Interviewer Shortcuts/Cheating
  - People will take path of least resistance when not properly supervised
    - Lazy sampling, friends/self completed etc...
    - Mostly after enough experience (boredom)



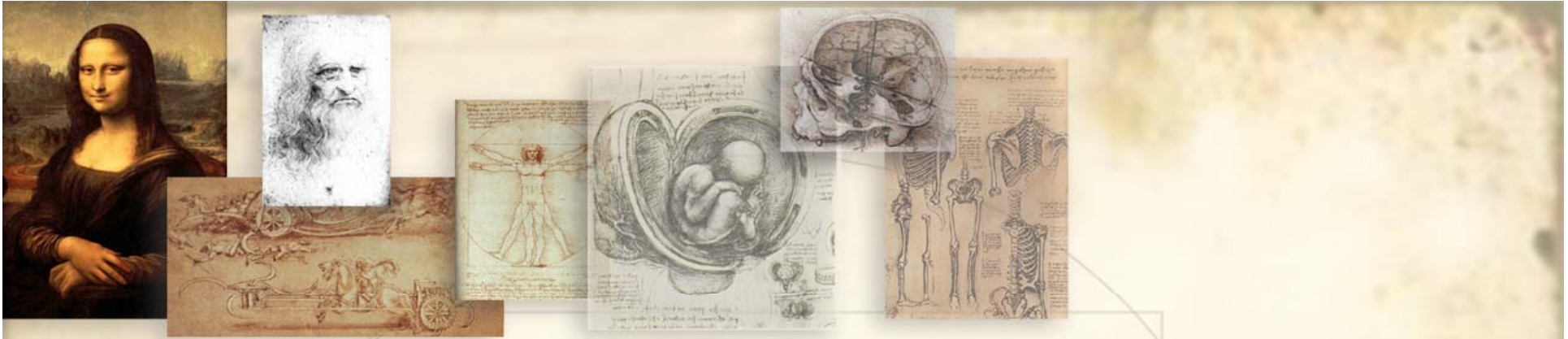
## Sampling – Abbreviated

- Sample should be representative of population of interest
- Larger the better
  - You can weight for lopsided representative
- Probability – Simple random usually best
  - Statistically more correct
  - Difficult to achieve
- Non-Probability
  - Cheap, easy and adjustable



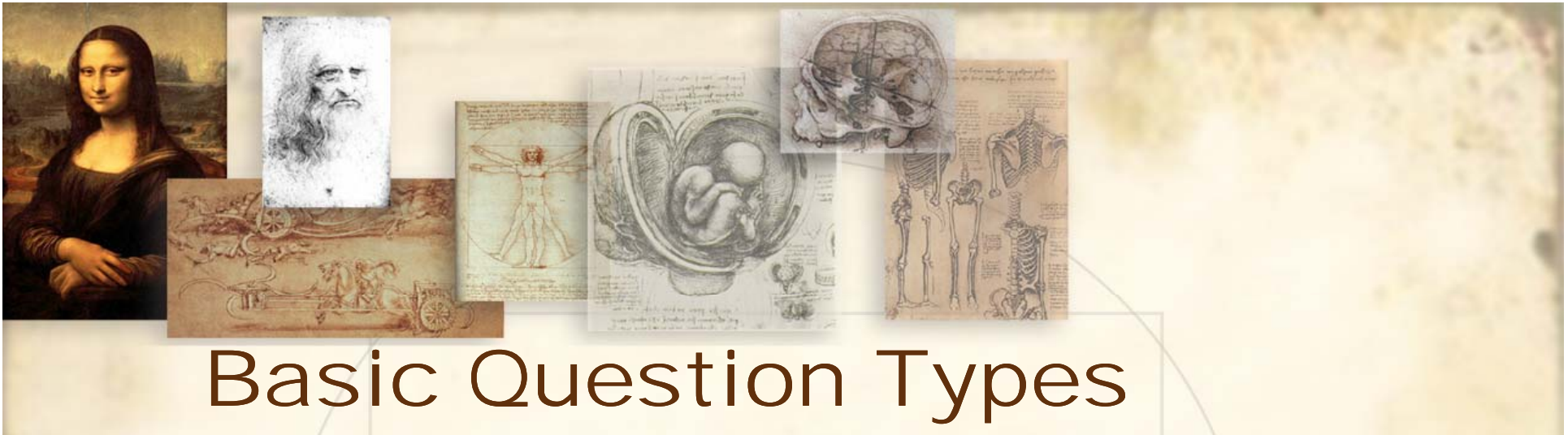
## Designing a Survey - Questions

- KISS – 12 yr old understands – unless expert interviews
- Scales – Ratio if possible (high as poss)
- Bad Questions
  - Unanswerable
  - Leading/Loaded
  - Double Barreled
  - Convoluted
  - Quick and Easy to answer



## Designing a Survey - Layout

- Appealing/Simple
- Whitespace/Crowded
- Flow
  - Instructions – thank you, you are important
  - First Question
    - Interesting
    - Answer is often “Yes”
  - More difficult or personal questions last



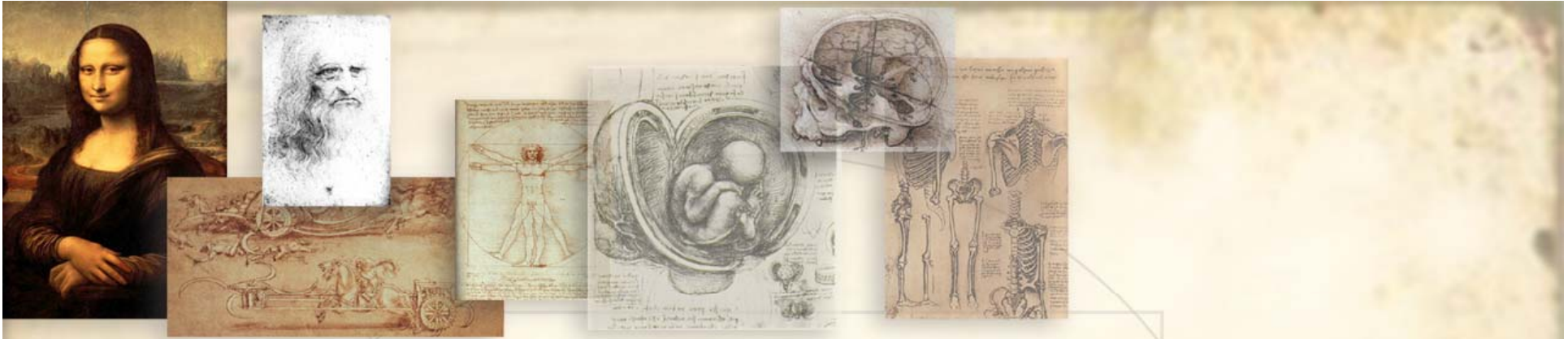
# Basic Question Types

Type	Advantages	Disadvantages	ROT
Semantic Diff Likert, Scale	Simple easy and quick	What does a "4" mean?	7 +/-2 points More the better
Ranking	Relative Importance	Space between ranks	No more than 6
Binary	Very Easy/Quick	Limited Info	0 not present, 1 present
Categories	Easier than Open ended	Precision	Use "other"
Constant Sum	Relative and weighted	Difficult for Avg. person to answer	Adjust to scale
All that Apply	Multiple category	Limited to membership	Code binary, use "other"



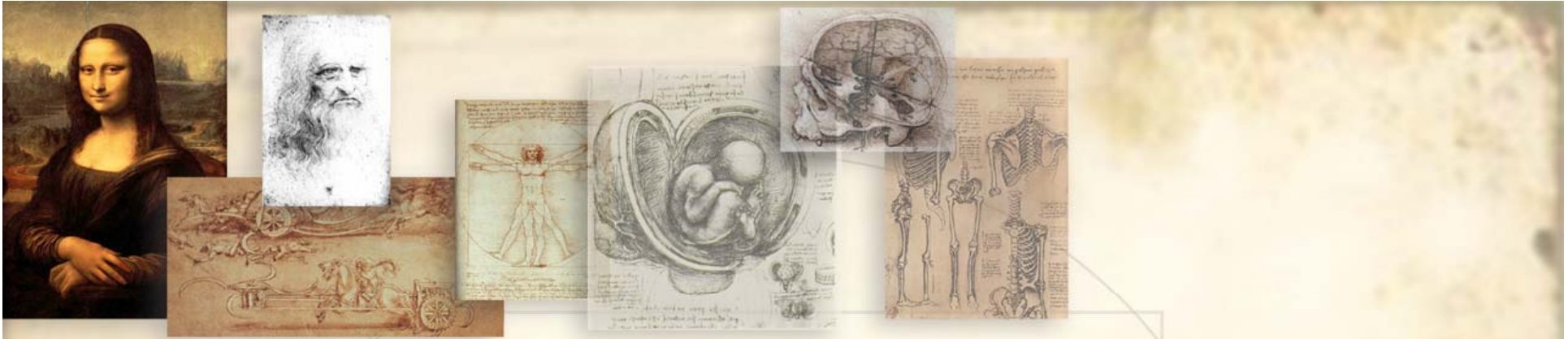
## Other Important Info

- Ch 9 does an excellent job on the process of survey design
- Exhibits 9.4 – 9.7 – good rules



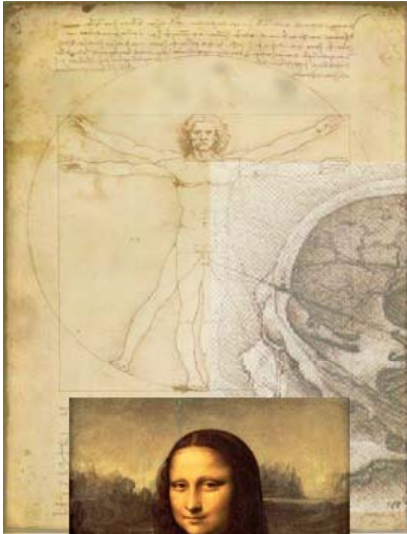
## Avoiding Errors

- Pretest, Pretest, Pretest
- Mutually exclusive, totally exhaustive
  - Other \_\_\_\_\_
  - Check all that apply
- CYA – client approval
  - e.g., Union Colony – billboards
- Open/Closed ended Q –
  - pay now or pay later



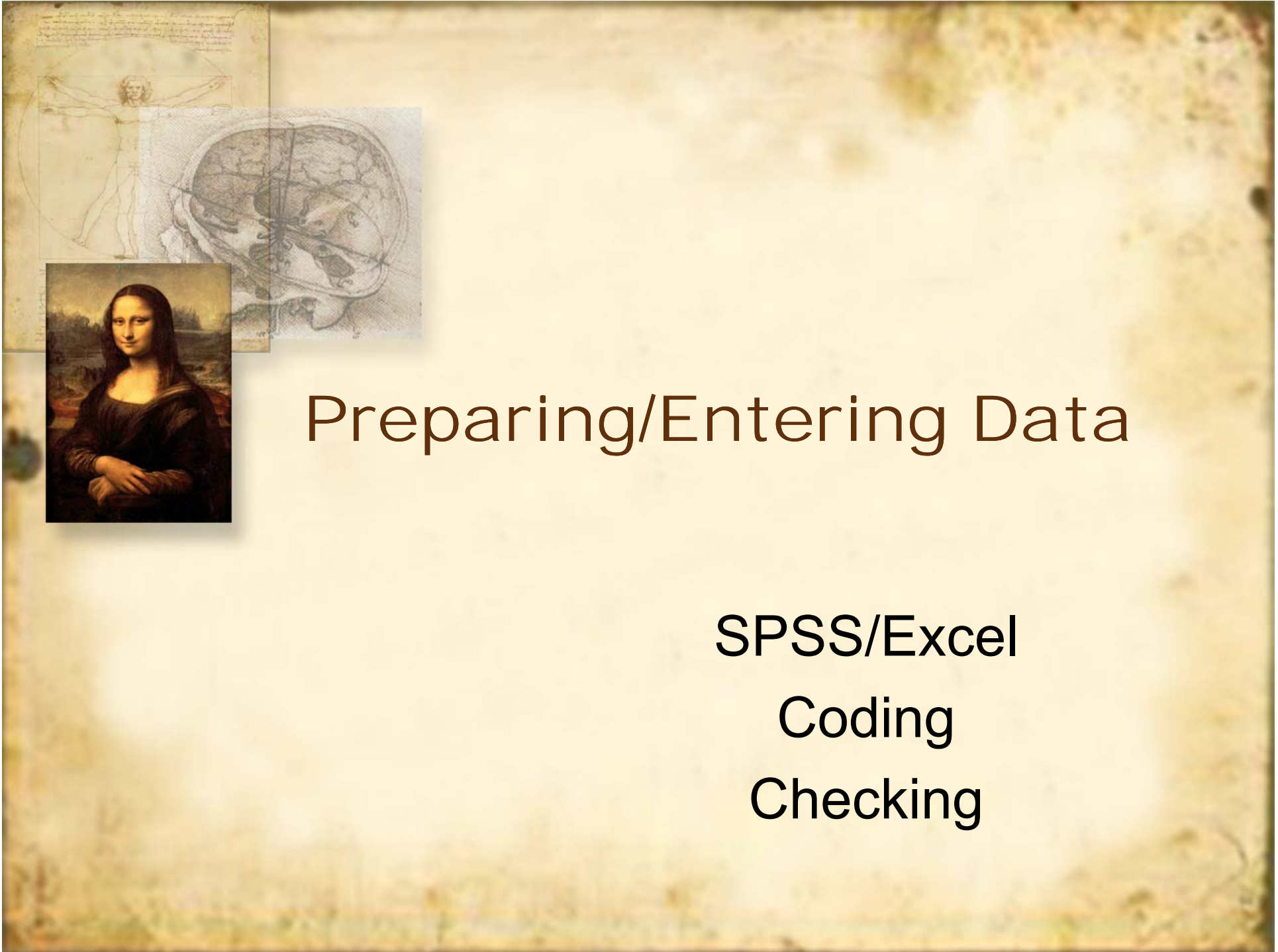
## Assignment for Wed 19-3-08 Part II

- Read assigned Chapters/Parts
- Create a “data needs” sheet
- Design a Survey/Data Collection Instrument



# Applied Marketing Research and Analysis Wednesday 19-3-08

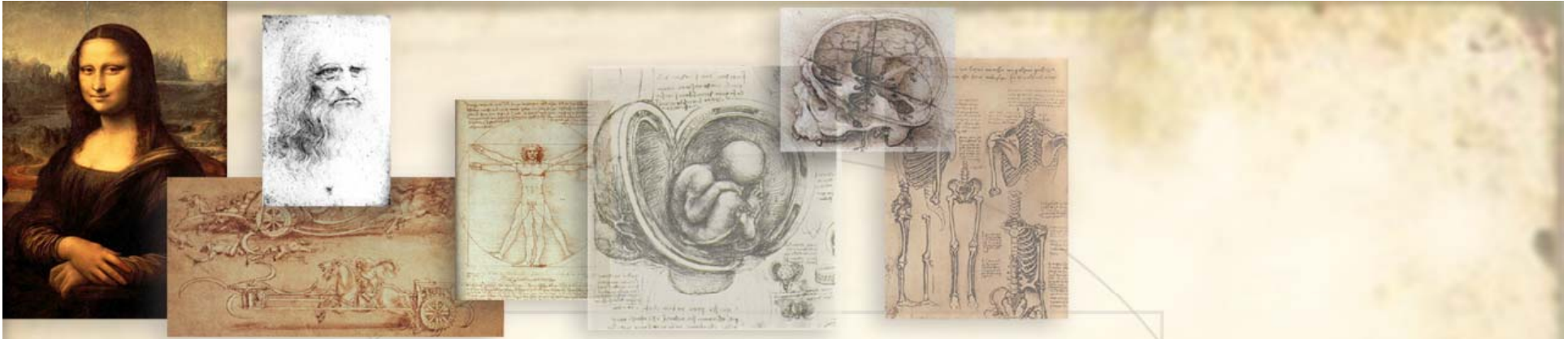
Data Analysis  
Interpretation  
Presentation



# Preparing/Entering Data

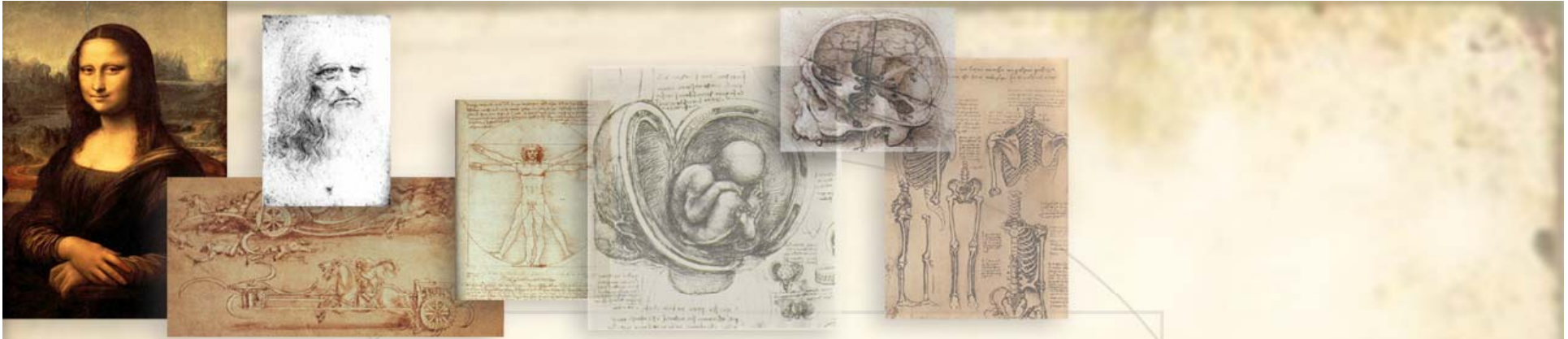
SPSS/Excel  
Coding  
Checking





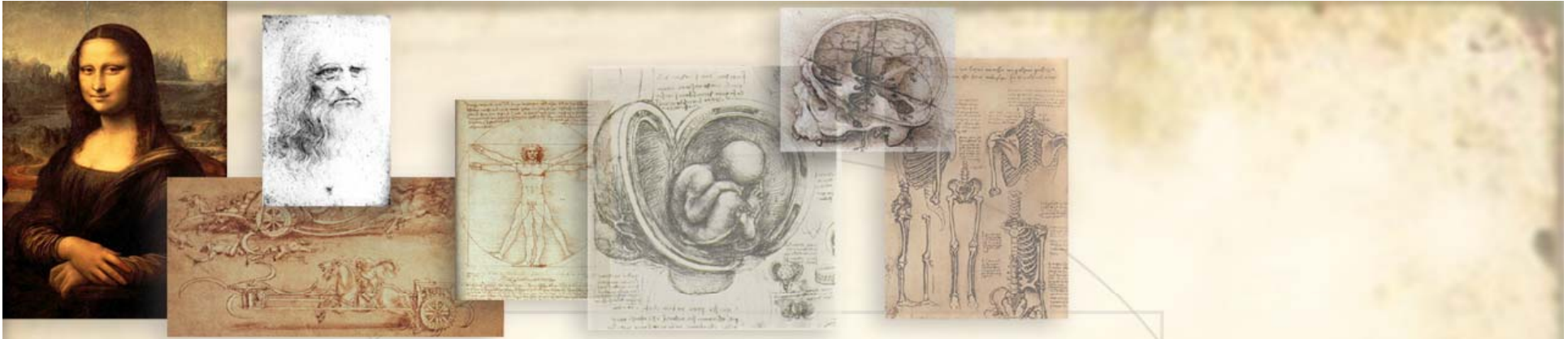
# SPSS

- Most can be done in Excel with more difficulty
- Windows
  - Data
    - Data View (tab)
    - Variable View (tab)
  - Output
  - Syntax – close unless repeating analysis



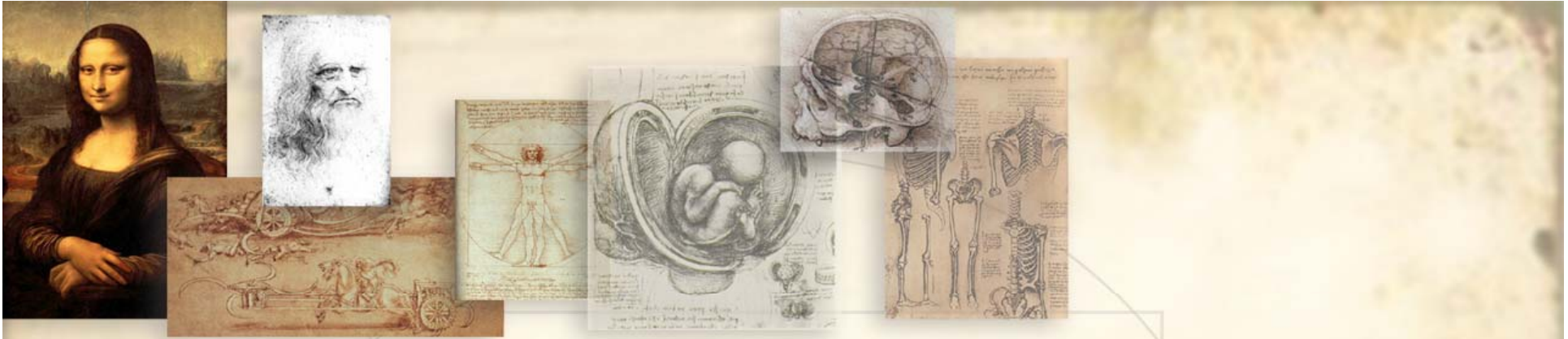
# Practical Coding

- Each survey has unique ID number
- Likert, Semantic Differential, Etc.
  - Enter number circled/measured
- Yes/No Questions
  - Dummy code – 1 yes, 0 if no
- Demographics
  - 1 if Male/Female (self), 0 if other
  - Mean of categories (central limit theorem)
- Check all that Apply
  - Separate variables for each choice
  - Dummy code – 1 if checked, 0 if not
- Leave empty if no answer



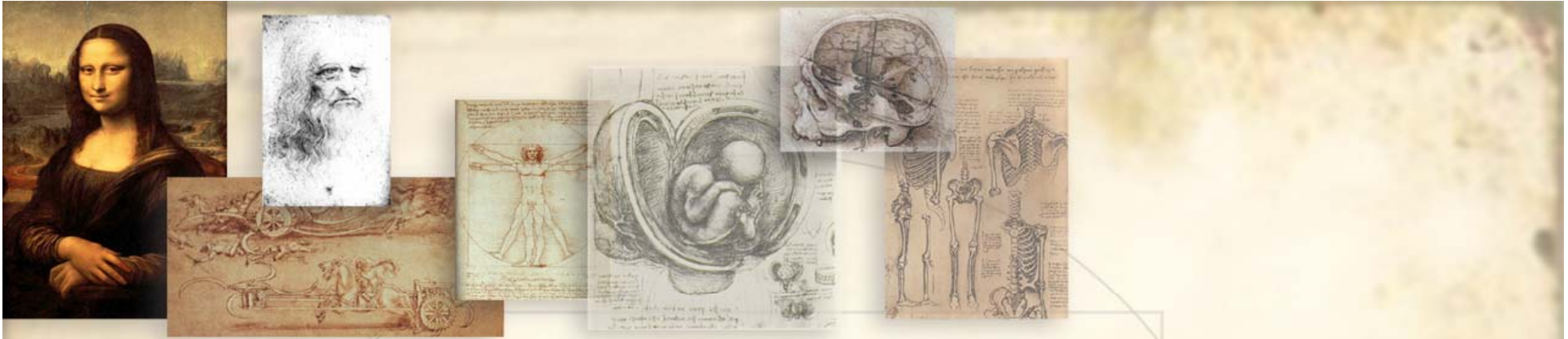
# Assignment

- Code your survey design from today's assignment



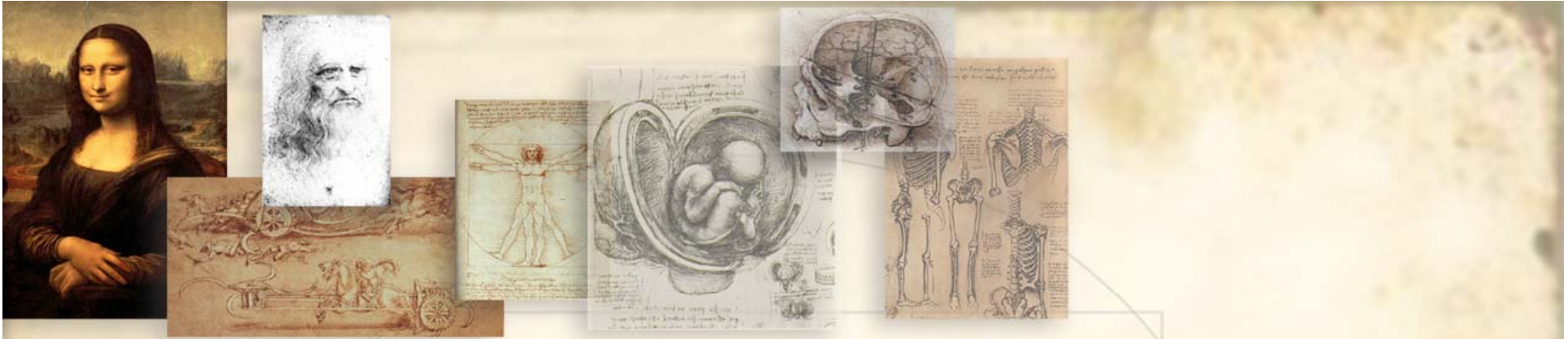
# Office Checking

- **Descriptives**
  - Check the range (min/max) and rest to see if they make sense
- **Frequencies**
  - Quick look to check for 'odd' entries
  - Do not "print" if interval level data



## Dataset Description

- **Subjects: 200+ Undergrad students from here**
- **Variables**
  - Music – Buy/Burn
  - Antecedents – as per dataset
  - Basic Demographics



## Functional Issues

- Creating New Variables
  - “Transform” -> “Compute”
- Selecting Cases (Subsets)
  - “Data” -> “Select Cases”
  - Remember to Unselect them when finished
- Split File
  - “Data” -> “Split File”
  - Remember to UnSplit file when finished



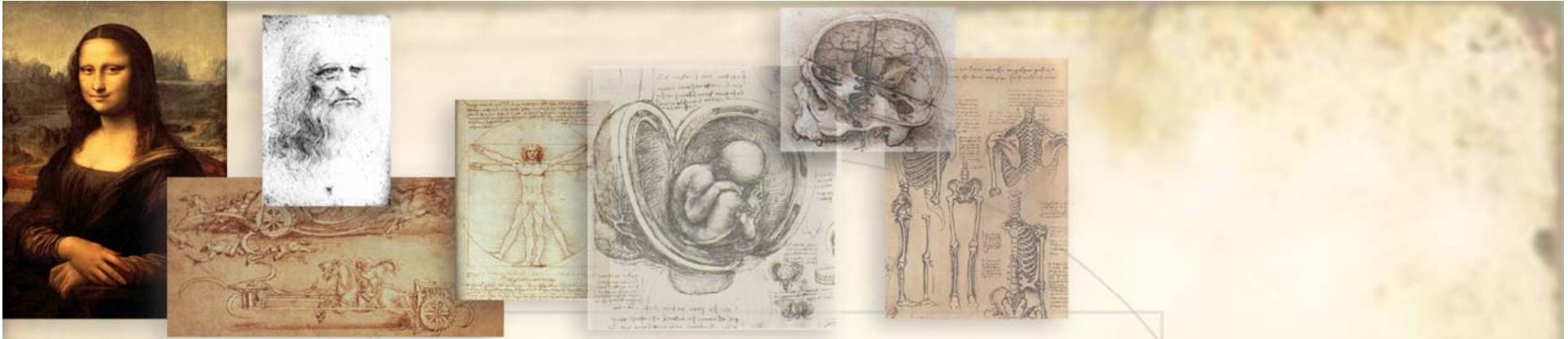
# Analysis and Interpretation

Crosstabs, t-tests, ANOVA, EFA

Means by

Correlation/Regression

Cluster Analysis



# General Uses of Statistics

- Descriptive
  - Basic look at data
- Data Manipulation
  - reformat data into alternative useable form
- Basic Examination
  - look at differences due to sampling
- Predictive/Causal
  - look at effects of one variable on another



# Comparing Occurrences

## Crosstabs

General Use: Descriptive/Examine

Main Use: Tables

Purpose: Tabulate Nominal Data

Interpretation: if  $SIG < 0.05$  then frequencies are not as expected

NOTE: Nominal Data (Frequencies)

		Income		Total
		Low	High	
Bought	Count	13	40	53
	% of Total	6.9%	21.2%	28.0%
Stole	Count	19	117	136
	% of Total	10.1%	61.9%	72.0%
Count		32	157	189
% of Total		16.9%	83.1%	100.0%



# Central Tendency - Means

'Means by' and t-test/ANOVA

Gen Use: Descriptive/Examine

Main Use: Describe differences

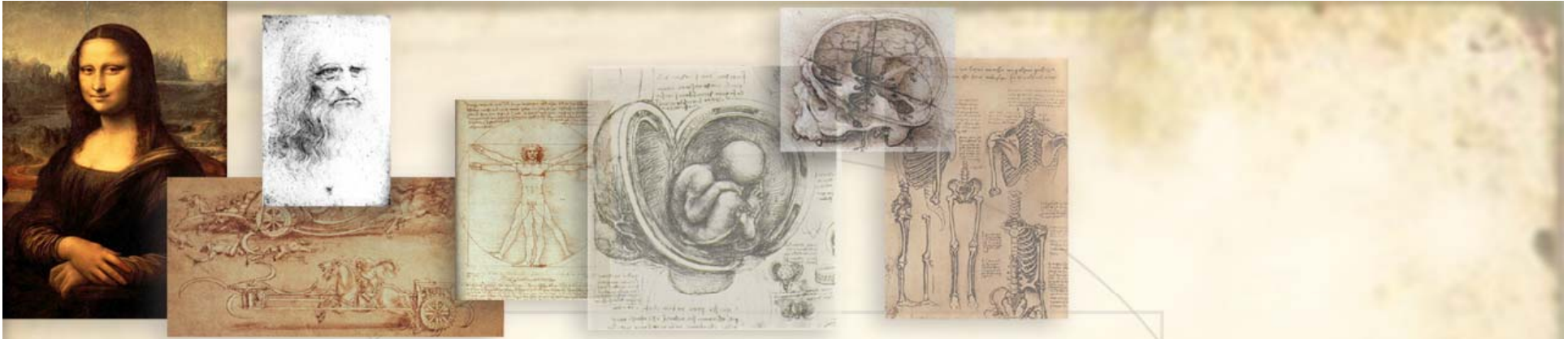
Purpose: Test if means are equal

Interpretation: SIG < 0.05 then means ARE significantly different

Remember means are Central Tendency – they do not describe all data -

Oven and Freezer Problems

Burn/ Buy	Mean	N	Std Dev
Buy	31.224	49	27.9
Burn	75.126	119	165.9
Total	62.321	168	147.7
T-value	2.79	Sig (p)	.006



# Data Reduction

## EFA – Factor Analysis

Gen Use: Data Manipulation

Main Use: Create Construct Measures

Purpose: Simplify related data

Interpretation: look at patterns of loadings above 0.50

Usually factor scores are saved and used for further analysis

Not used often in business

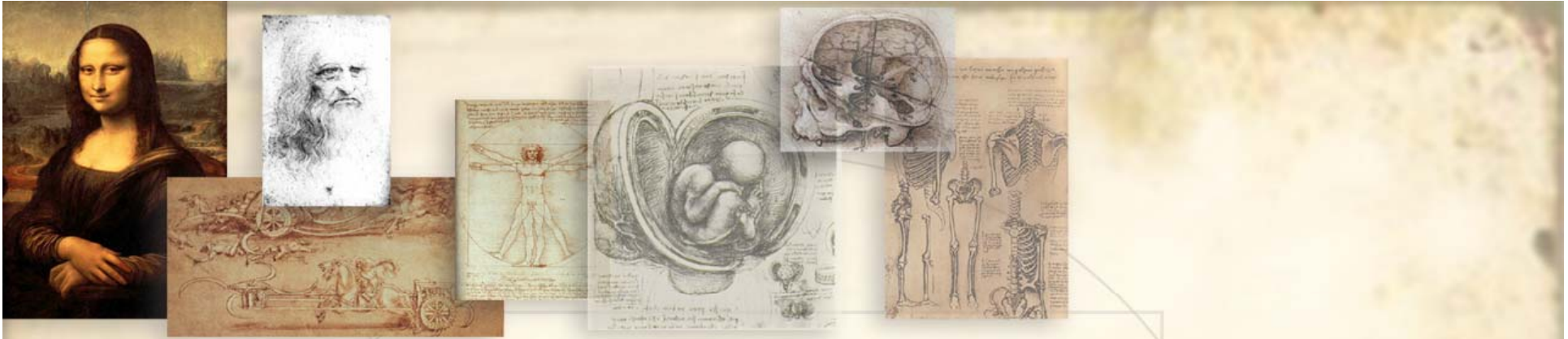
- Attitudes

- Formative

- Summative

- Reflective

- EFA – combine to a single score with Mean = 0,
- Std Dev = 1



## Simple EFA

- Analyze ->Data Reduction -> Factor
- Choose Variables
- “Rotation” – choose “varimax”
- “Scores” – Click box to “save as variable”
- “Options” – Last box
  - Suppress Absolute Values.... Put 0.5 in box



# EFA Example

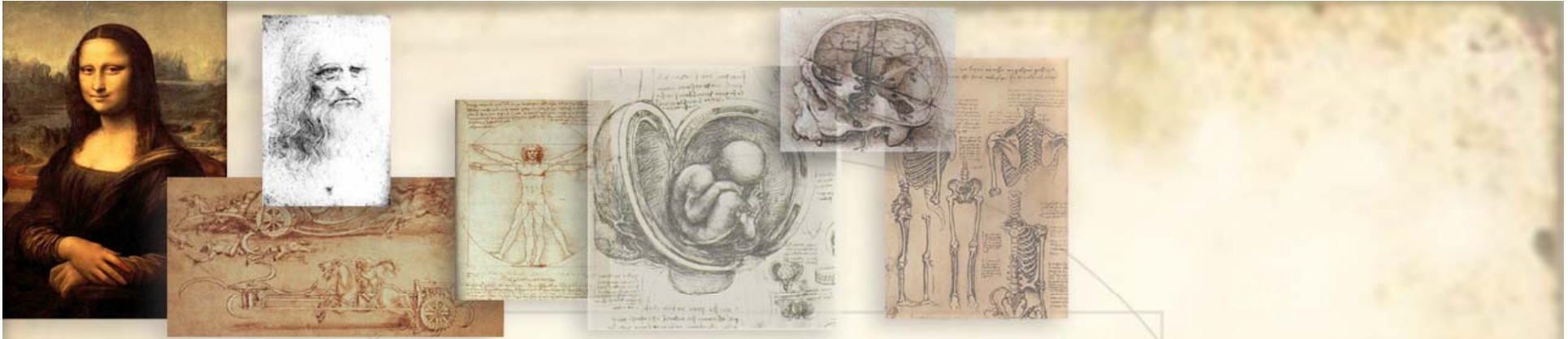
## Rotated Component Matrix(a)

	Component	
	1	2
CopyEth1	.834	
CopyEth2	.849	
CopyEth3	.882	
CopyRisk1		.588
CopyRisk2		.830
CopyRisk3		.723

Extraction Method: Principal Component Analysis.

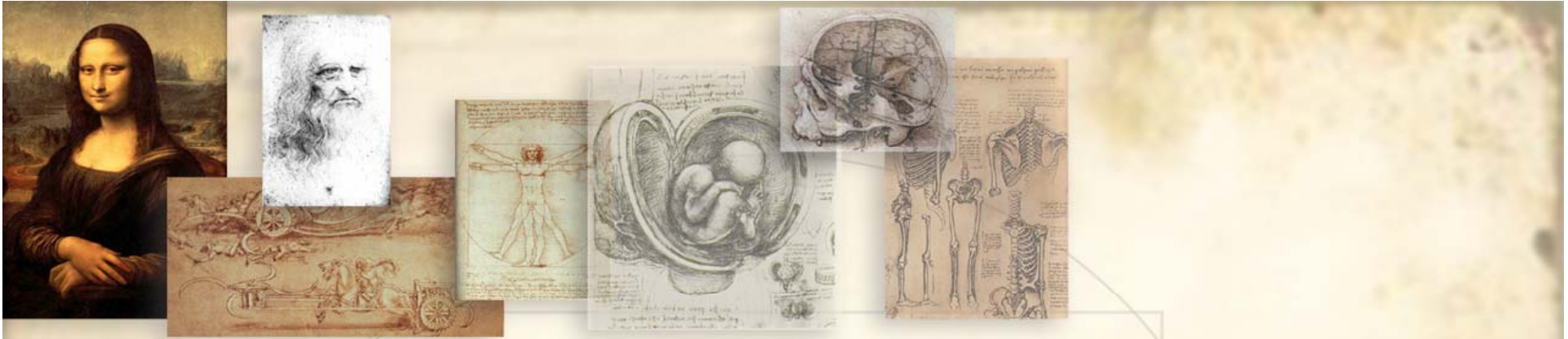
Rotation Method: Varimax with Kaiser Normalization.

a Rotation converged in 3 iterations.



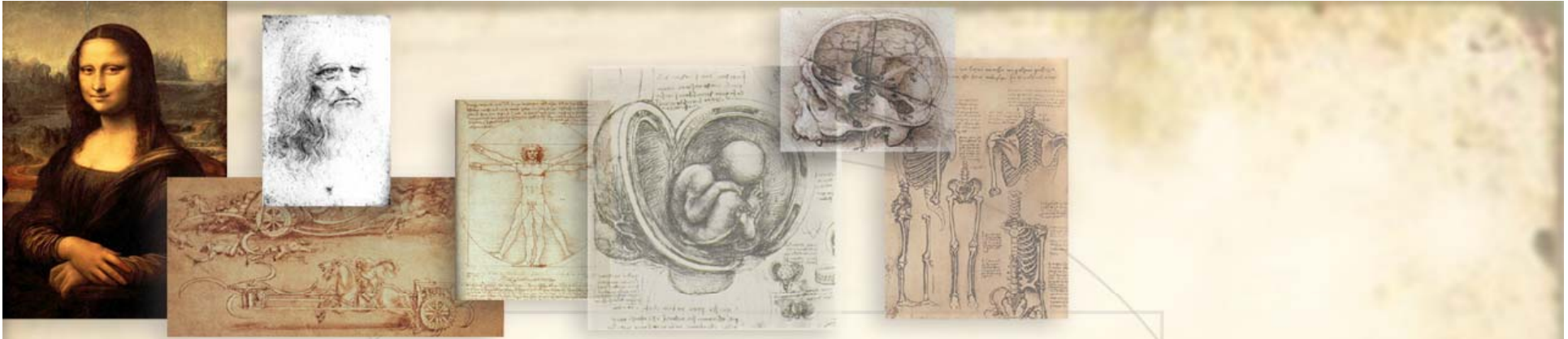
## Correlation

- Gen Use: Descriptive/Examine
- Main Use: Initial Inspection of data
- Purpose: Examine bivariate relationships
- Interpretation:  $SIG < 0.05$  then variables are related/correlated Exhibit 13.5 (p.287)
- Exhibits 13.1-13.4 (p. 283-286)



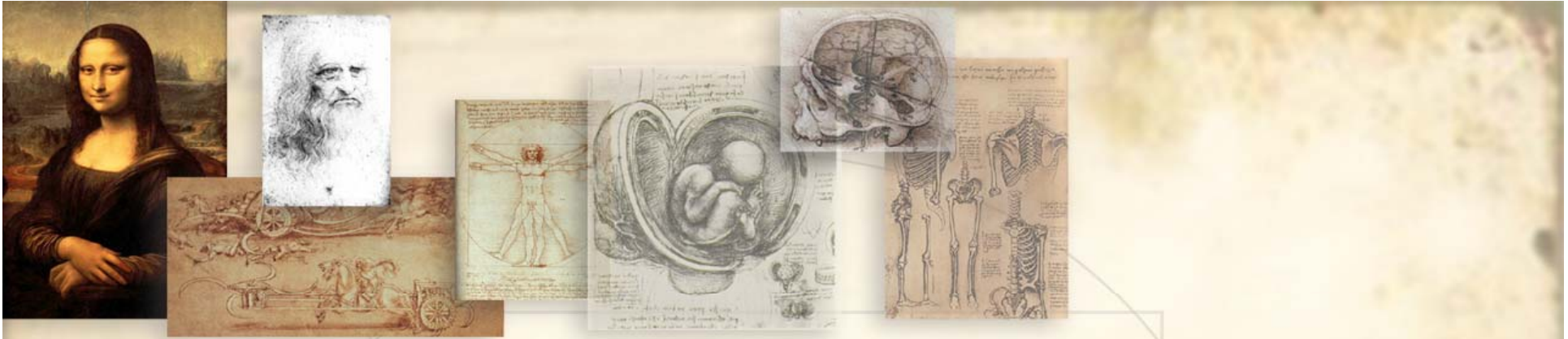
# Regression

- Main Use: Most used/useful/versatile of all analysis techniques
- Purpose: Show relationships/effects
  - Draws a line
- Basic Interpretation
  - ANOVA F – if  $SIG < 0.05$  -> Model good
  - R-Squared - % of variance explained
  - Coef t – if  $SIG < 0.05$  then Coef  $\neq 0$  (good)



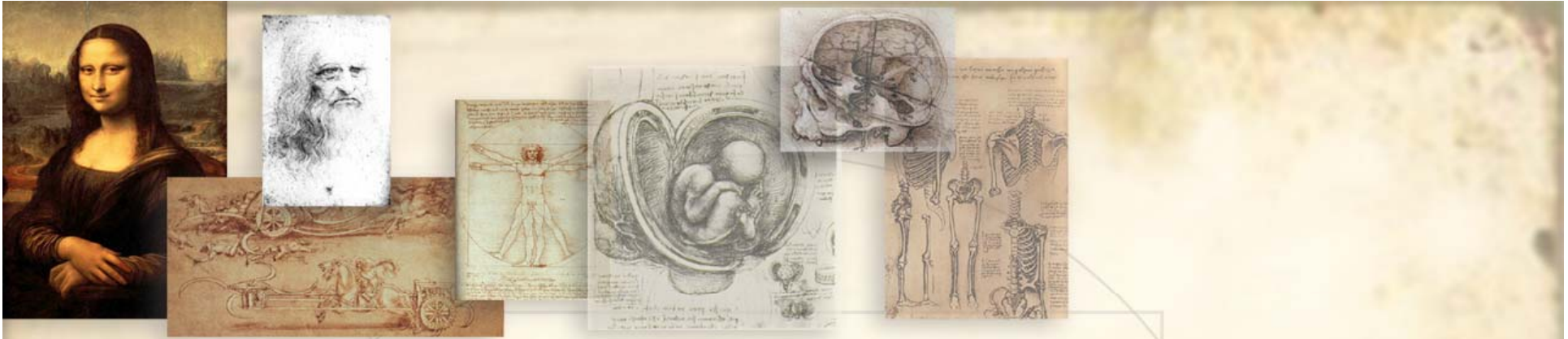
## Coefficient Interpretation

- Y Intercept – basic starting point assuming all X variables equal 0
- For every unit change in X there is (on average) a (coef) change in Y
- Dummy Variables – for a change in X from 0 to 1, there is a one time (coef) change in Y



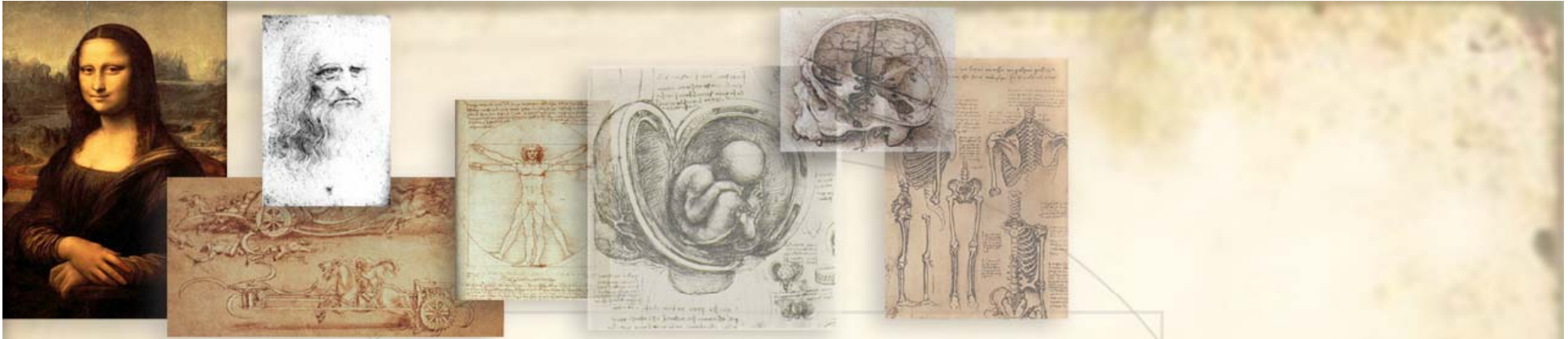
## e.g., Regression Interpretation

- if F/Anova is significant ( $\text{sig} < 0.05$ ) then continue with interpretation
- Look at R-Squared
  - Percent of Y explained by X variables
- Individual Coefficients
  - First (B) and last (Sig) number columns
  - Look at Sig – if  $< 0.05$  then B is not 0 (ie., variable has an effect)
  - Interpret B as degree of change in Dependent(Y) for every unit change in Independent (X)



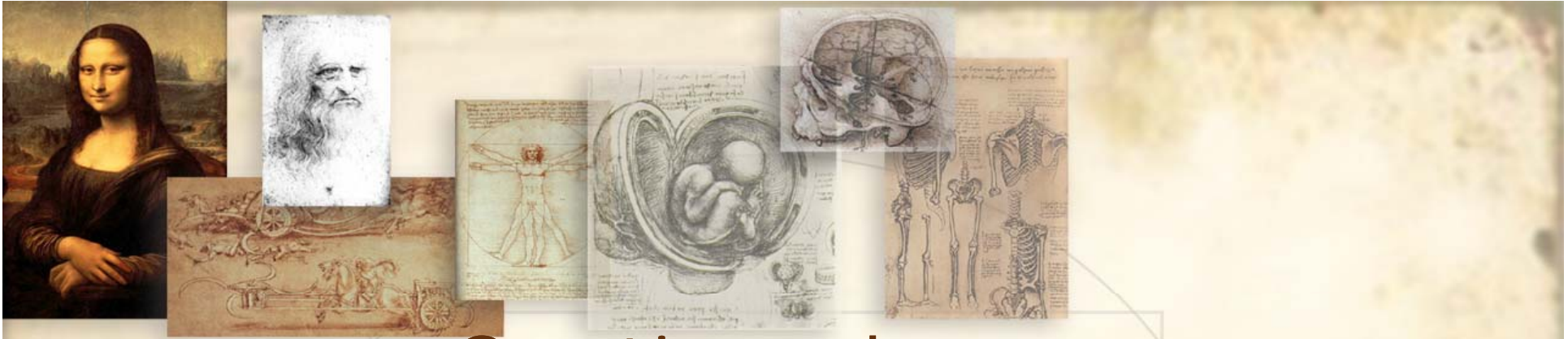
# Intro to Cluster Analysis

- Main Use: Segmentation
- Purpose: Form homogenous groups
- Interpretation: Based on the centers of each group – Normative identification
- Usually results are used for further analysis – ‘save group membership’



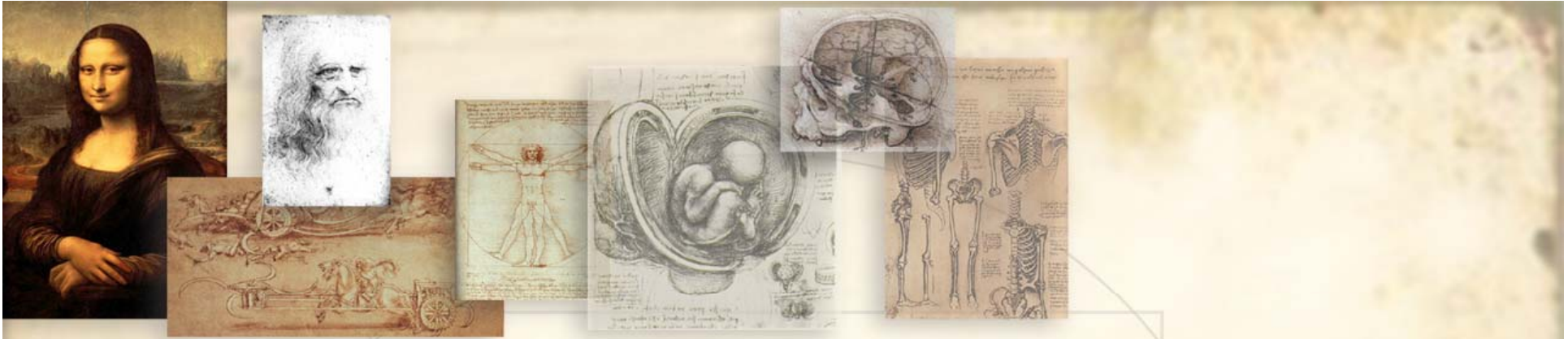
## Two types of Cluster

- Menu – Analyze -> Classify ->
- Two Step Cluster
  - Plug and Play
    - Put in variables in appropriate spot
    - Ratio and Interval in “Continuous”
    - “output” – click
      - “create cluster membership variable” box
  - Number of clusters determined by data



## Continued

- **K – means cluster**
  - Typically only continuous variables – categorical variables will normally result in separate clusters
  - User chooses number of clusters
    - Either using Logic
    - Or more often – trying different solution to see what works best
      - Best – no “small” clusters (few members)
      - All clusters are interpretable



# Final Assignment

- Write one intelligent sentence explaining the results

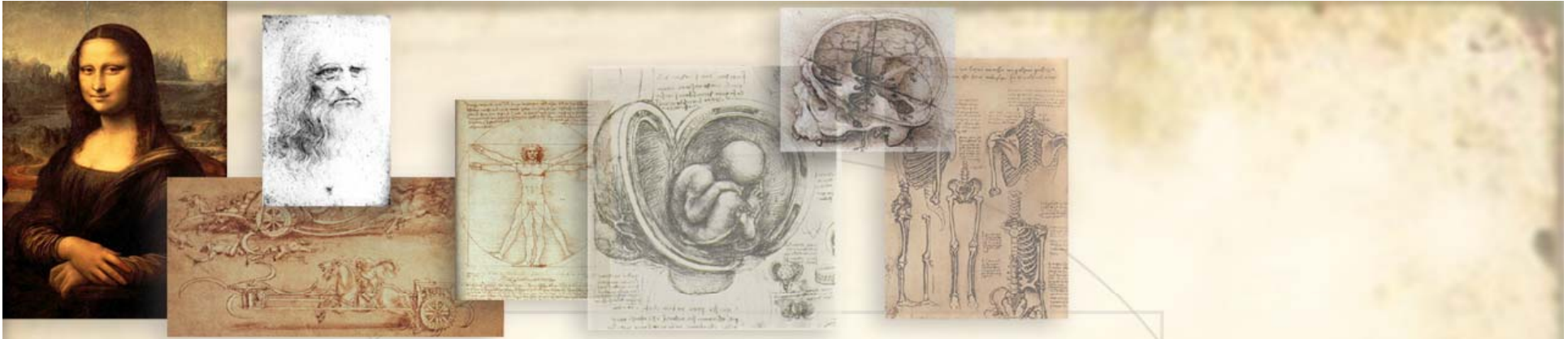


# Presentation

Educational vs Business

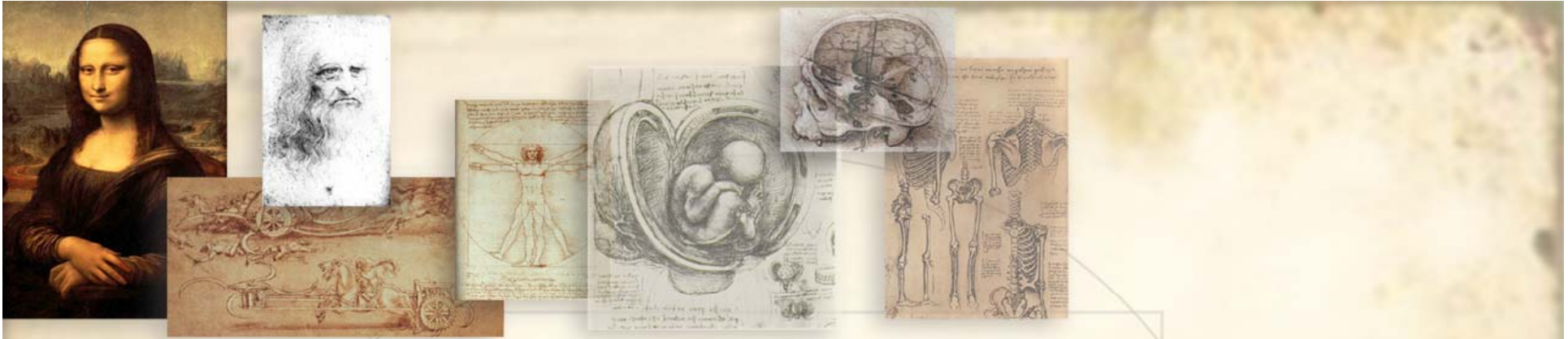
KISS Principle

Improving Presentations ROT



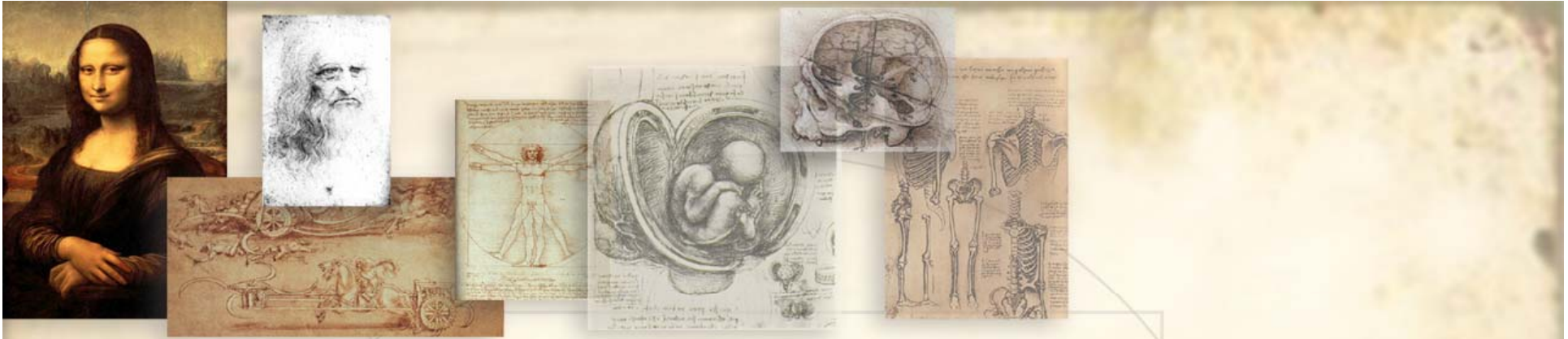
# KISS

- Keep it Simple for the Stupid
  - Executive Summary
  - Graphs/Tables/Exhibits
  - No “stat talk”
    - Exception – technical audience
  - Brief, concise and to the point



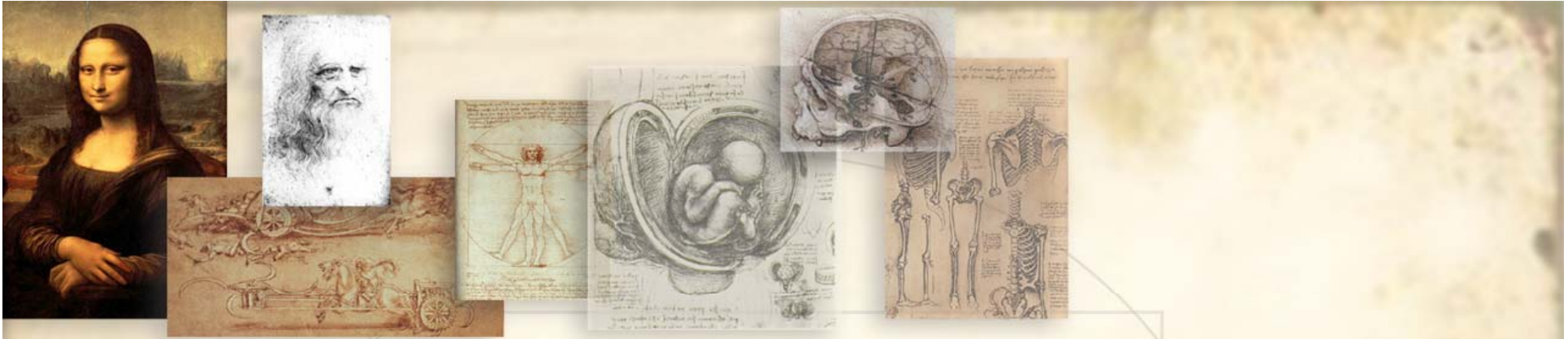
# Rules of Thumb for Presentation

- Rule of affinity
  - Art/graphics/interest – a good visual on every page
- Know your audience
  - Diverse audience is hardest to present to
  - Information at appropriate level
  - When in doubt – put it in an appendix
- Backup slides/handouts with more/additional information – anticipate questions



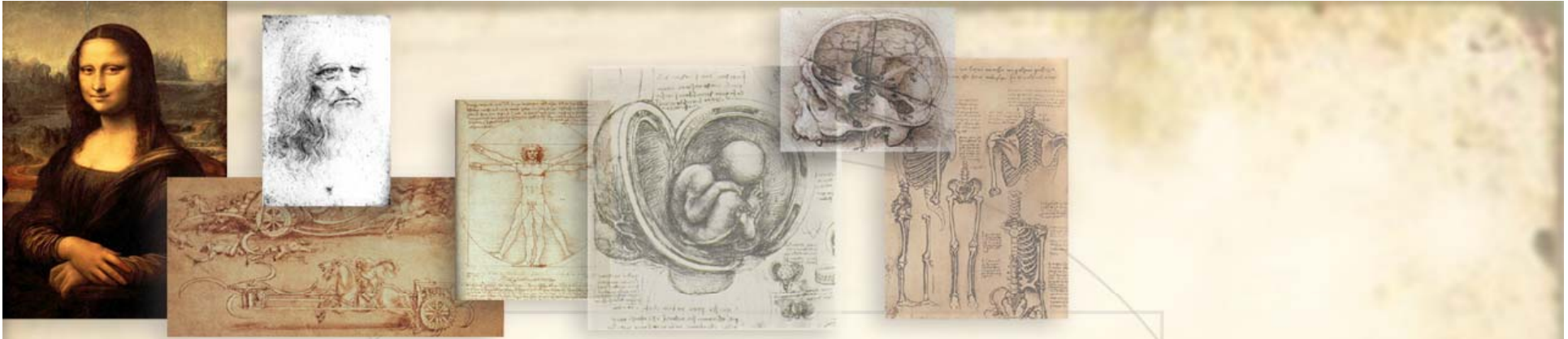
# Written Presentations

- **Executive Summary**
  - Summarizes – does not introduce
  - ONE page
  - All important findings contained
  - Very limited details
  - For the executive with limited time
  - Ask yourself – “if I had 2 minutes in his office, what would I tell him about this”



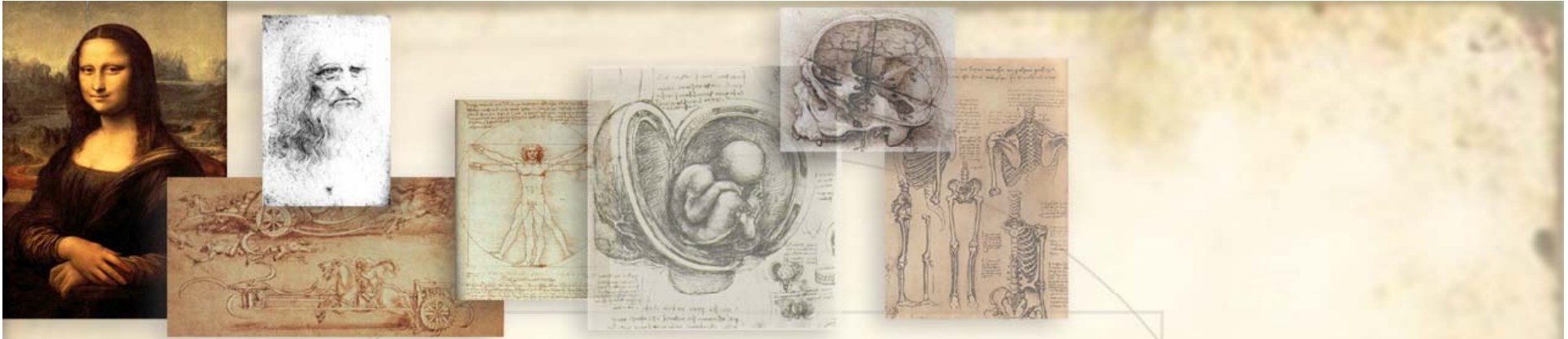
# Written Presentations

- **Presentation Style**
  - Tell a story – relate and transition
  - Concise, to the point – the shorter it is, the more likely it will be read
  - **Active Voice**
    - No - I will research this issue
    - Yes – This issue needs researched
  - If research – keep opinions, judgments to the conclusion/summary
  - **SHOW ME.....**



# Oral Presentation/Powerpoint

- Backups – power fails, computers corrupt
- Picture worth 1000 words
- Very limited text, bullets (not like edu)
- Glance test – get the general point
- Don't assume – interpret for them
- Details are Oral – graphics are for visuals
- Any less than 20pt text is unreadable
- Duplicate handouts ensure they won't pay attn to your presentation – handouts should be complimentary to the material in the presentation



# Random thoughts in Parting