



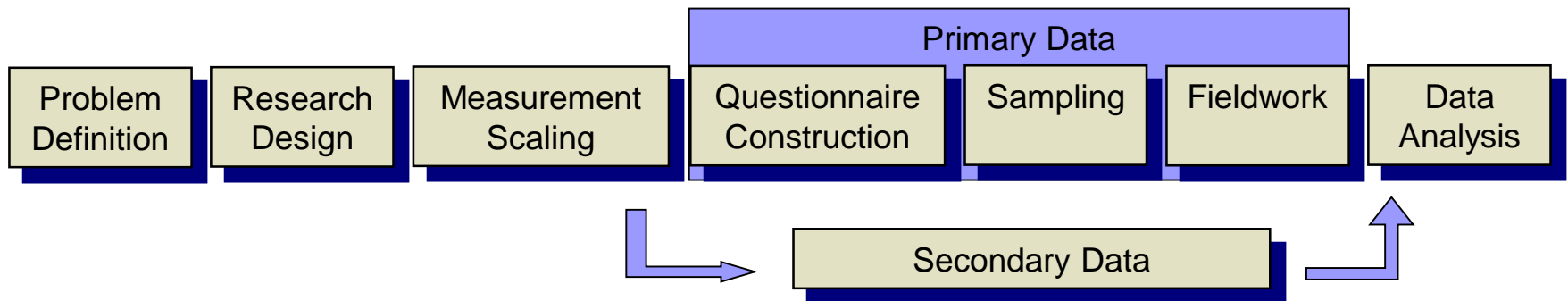
# Cross-National Research and Data Analysis

Session 7



# Agenda

- Marketing Research Process:



# Fraction of Customers Having Churned over the Last 6 Months (2007)

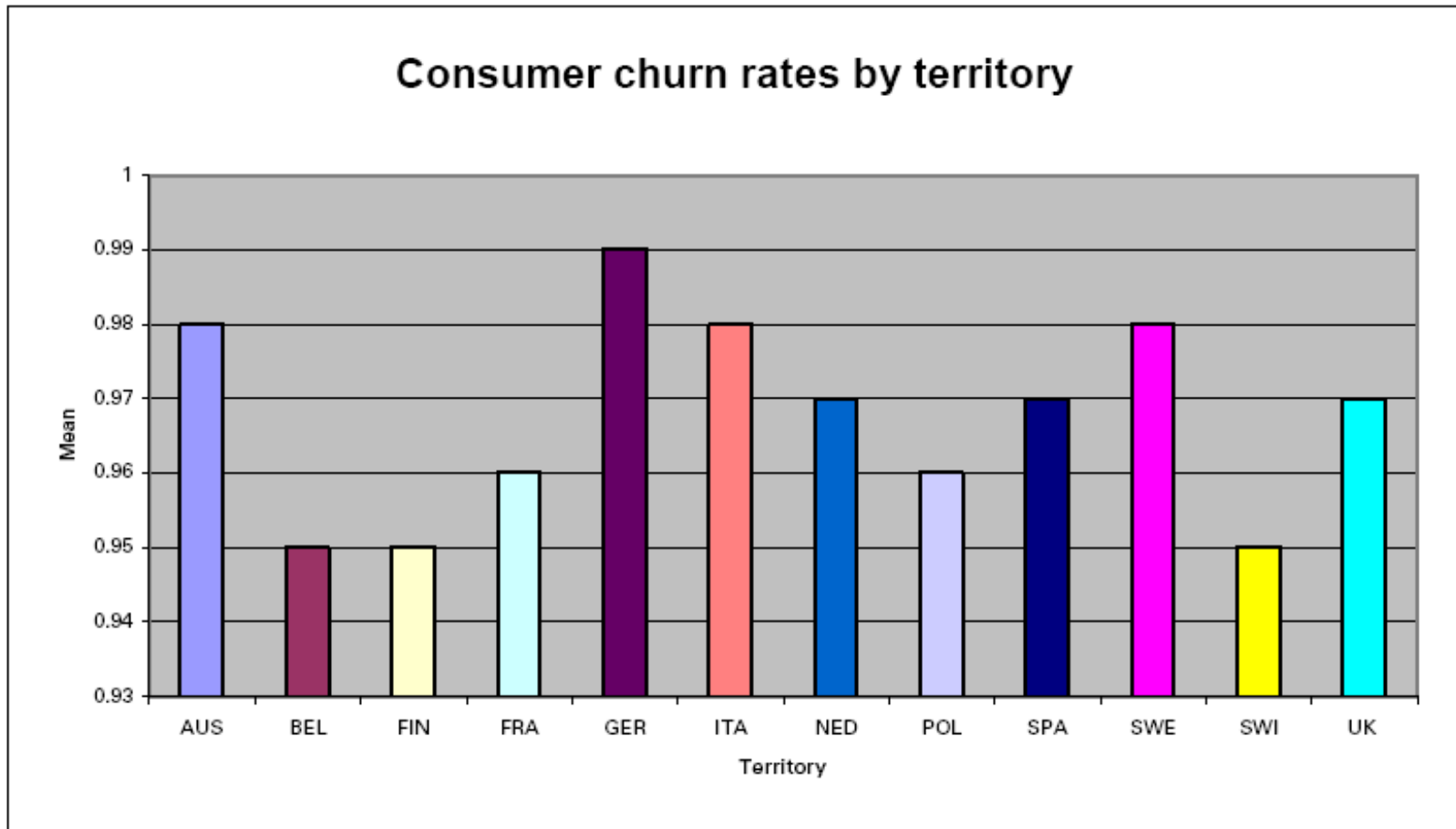


Figure 1. Consumer churn rates by territory

# Fraction of Customers Having Churned by Industry (Europe - 2007)

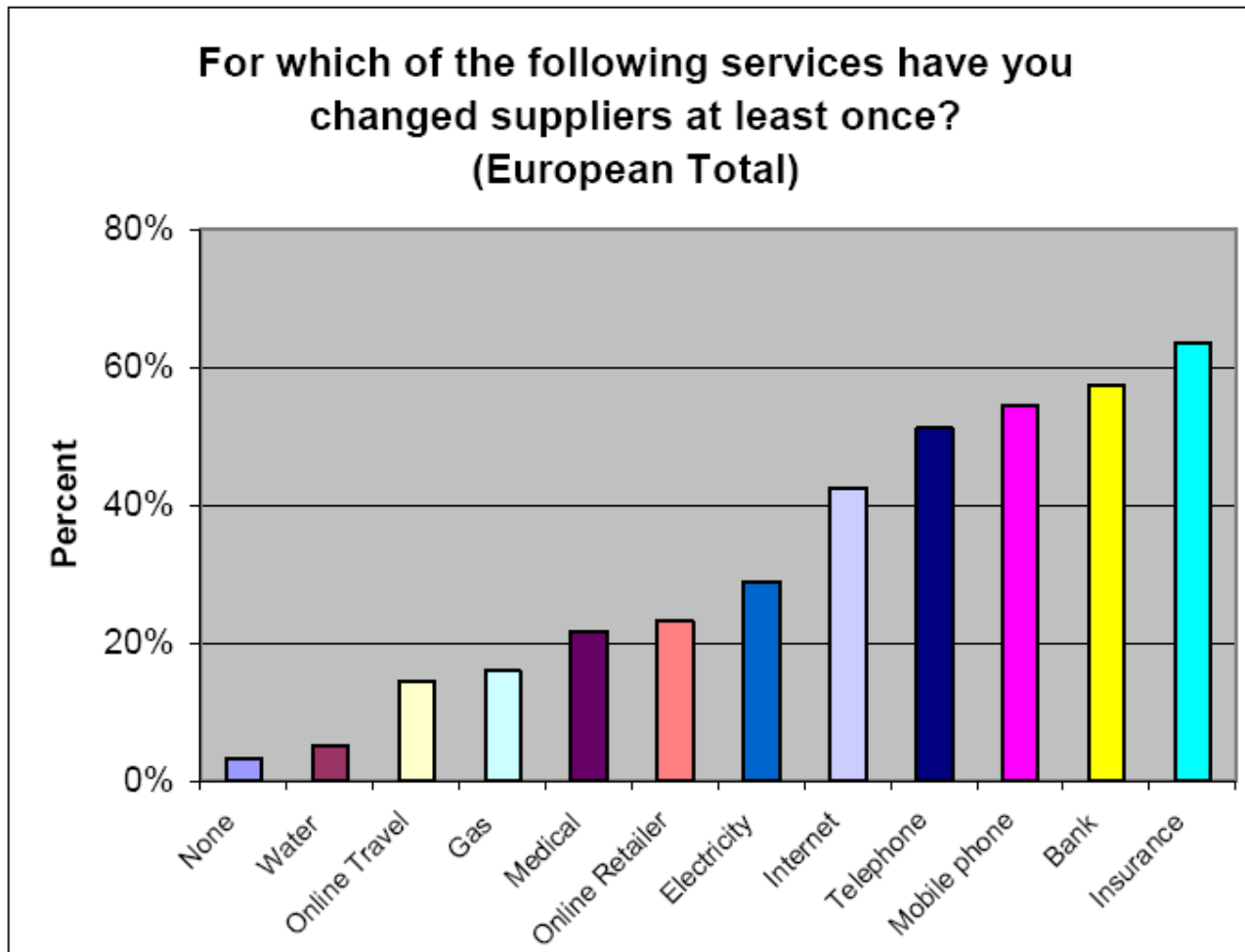
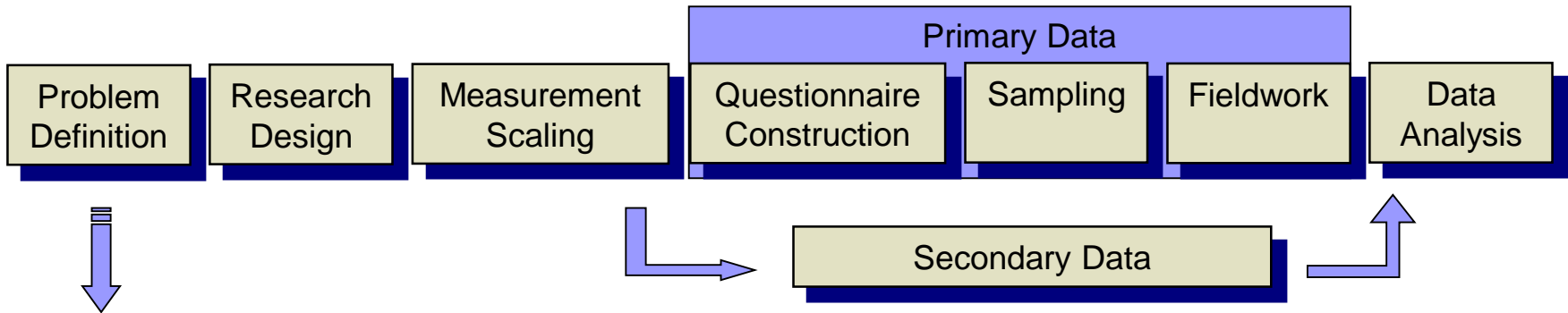


Figure 2. Supplied services changed at least once (European total)

# Marketing Research Process



Decision Problem

Example: Decide on action w.r.t. customer attrition in Europe

- Management level:
  - Corporate
  - Regional
  - Local
- Decision type
  - Strategic: market entry, segmentation, positioning
  - Tactical: marketing mix decisions based on the strategic options

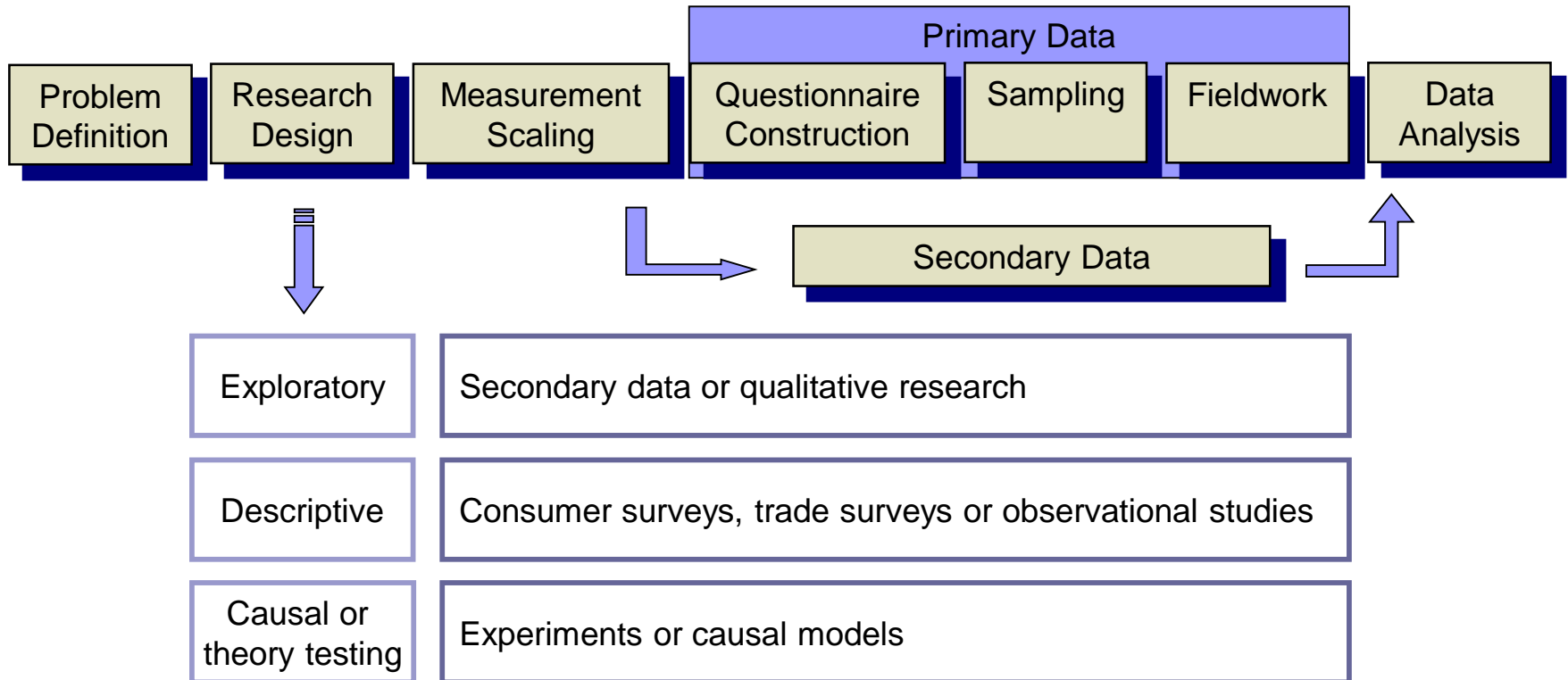
Research Problem

Example: What are the drivers of (causes for) customer attrition?

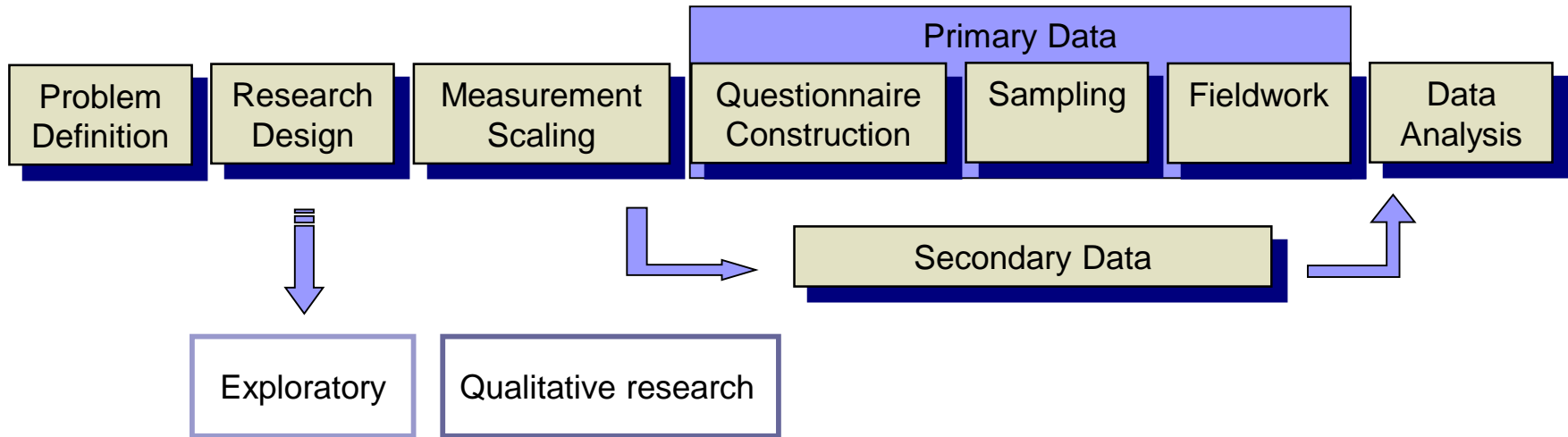
Unit of analysis

World, regions, trade blocs, countries, cities, ... → within-country homogeneity ?

# Marketing Research Process



# Marketing Research Process



- Focus groups

- 8 to 10 consumers and a moderator
- Video, one-way mirror
- Mainly for product policy decisions (ideas for new products)
- PROS: Inexpensive, quick
- CONS: Unrepresentative sample



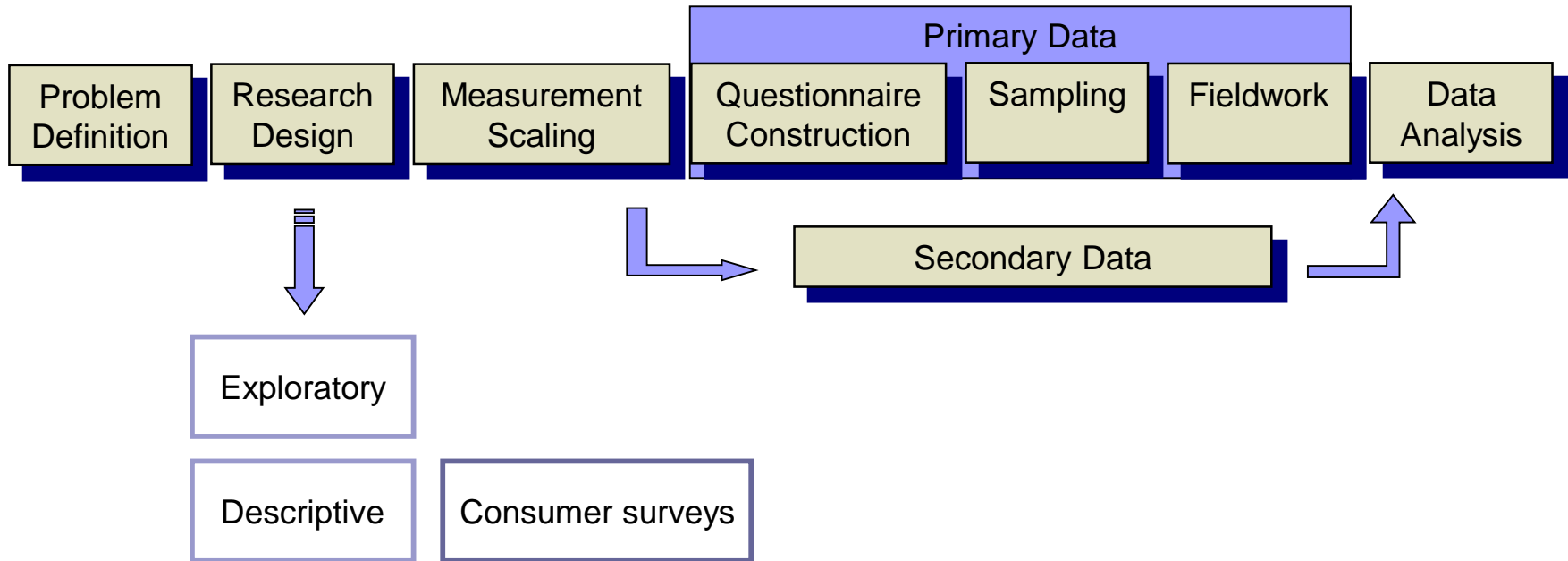
# An Example: Coke Light – Coke Zero

- Who still remembers the Coca-Cola-commercial on the 11.30 am arrival in the office of the coke light delivery guy?
  - Aim: Coke Light is trendy, sexy and refreshing
- Focus Group
  - Coke is associated with sexiness but as side effect of the commercial, also with a feminine consumption of the brand (few male consumers)
- Solution found: Coke Zero





# Marketing Research Process



- Larger groups (> 500)
- Random samples
- Via mail, phone, e-mail, internet, in person
- Mainly for ads (pretest, post test, recall scores, media habits), promotion (responses to promotion), segmentation and positioning (consumer habits, product benefits), satisfaction measurement, competition evaluation

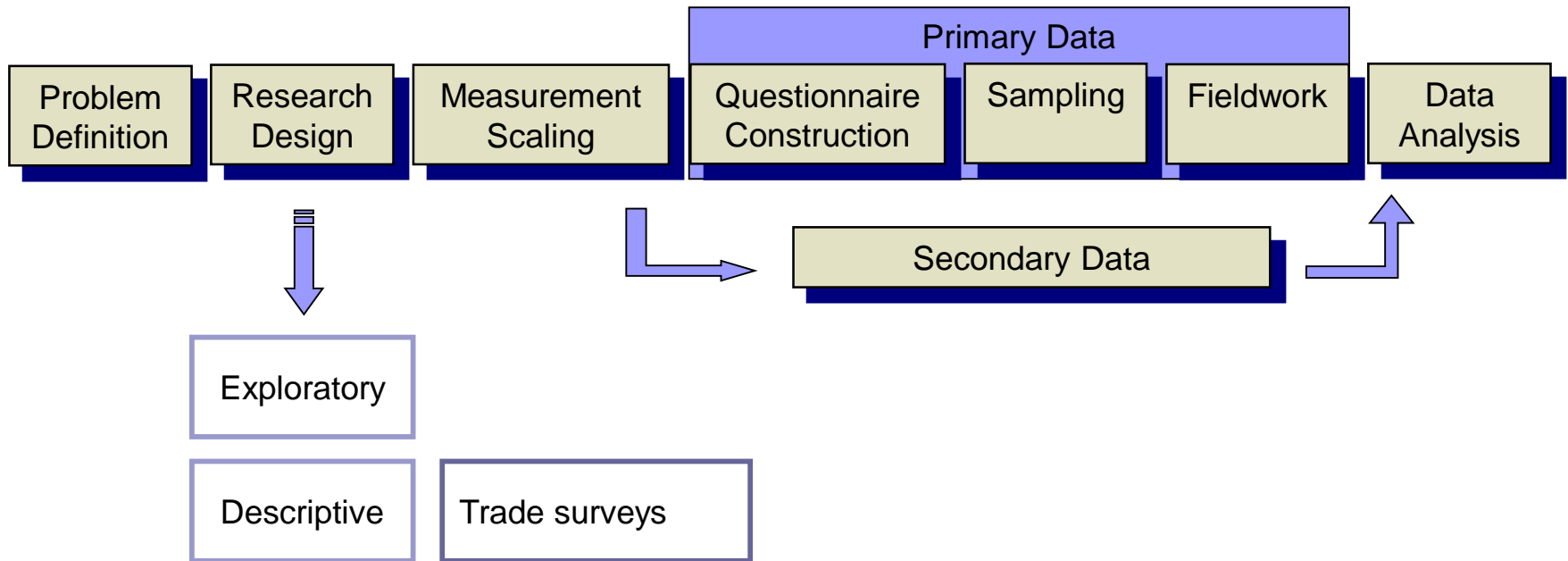
# Consumer Survey: Car Brand Perceptions

- *Despite experiencing one of the most turbulent years ever in the automotive market, Chevrolet, Ford, and Subaru have benefited from a dramatic increase in overall brand perception among car owners. According to Consumer Reports 2010 Car Brand Perception Survey, all three saw at least a 30-point increase in their overall perception scores over last year's scores. The scores reflect how consumers perceive each brand in seven categories: Safety, quality, value, performance, design/style, technology/innovation, and environmentally friendly/green.*

*Toyota maintained its Number one spot. But overall score improvements helped Ford to slip past Honda into second place and allowed Chevrolet to rise to fourth place, from ninth. Subaru's 39-point increase, which was the largest of any brand and almost double its 2009 score, boosted the brand to ninth place, from last year's 18th position*

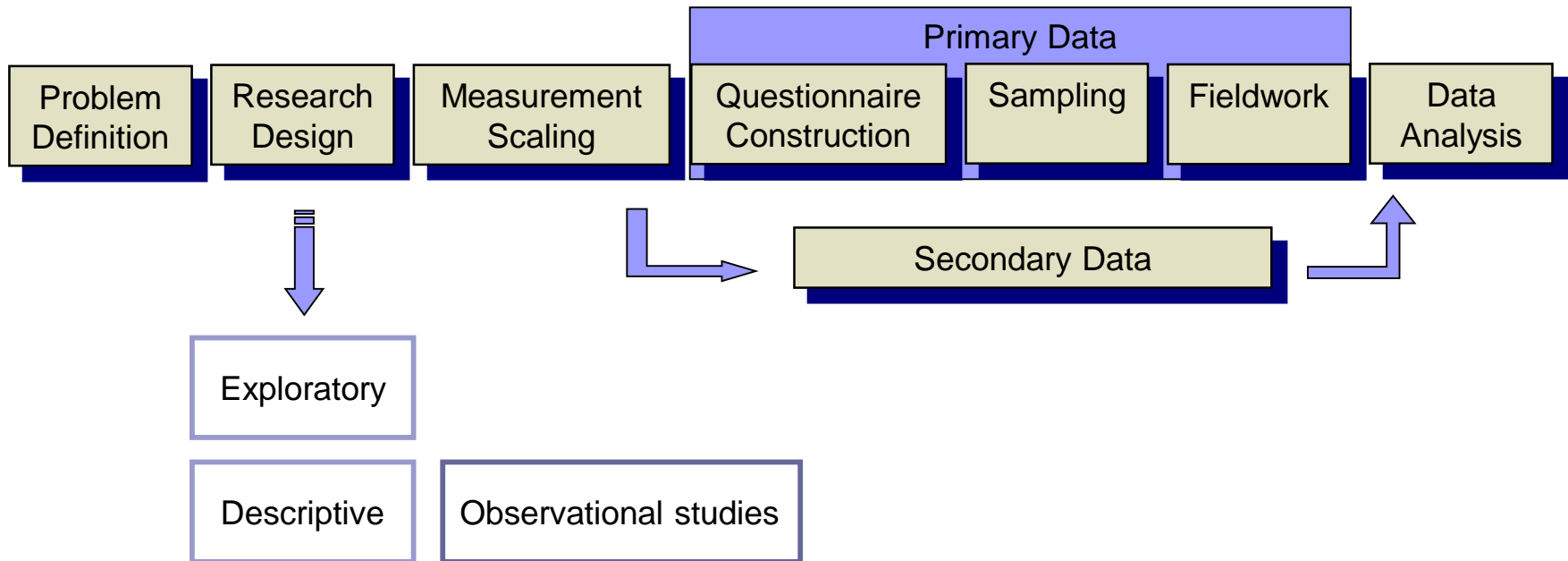
- Source: [2010 Consumer Reports Car Brand Perceptions Survey](#)

# Marketing Research Process



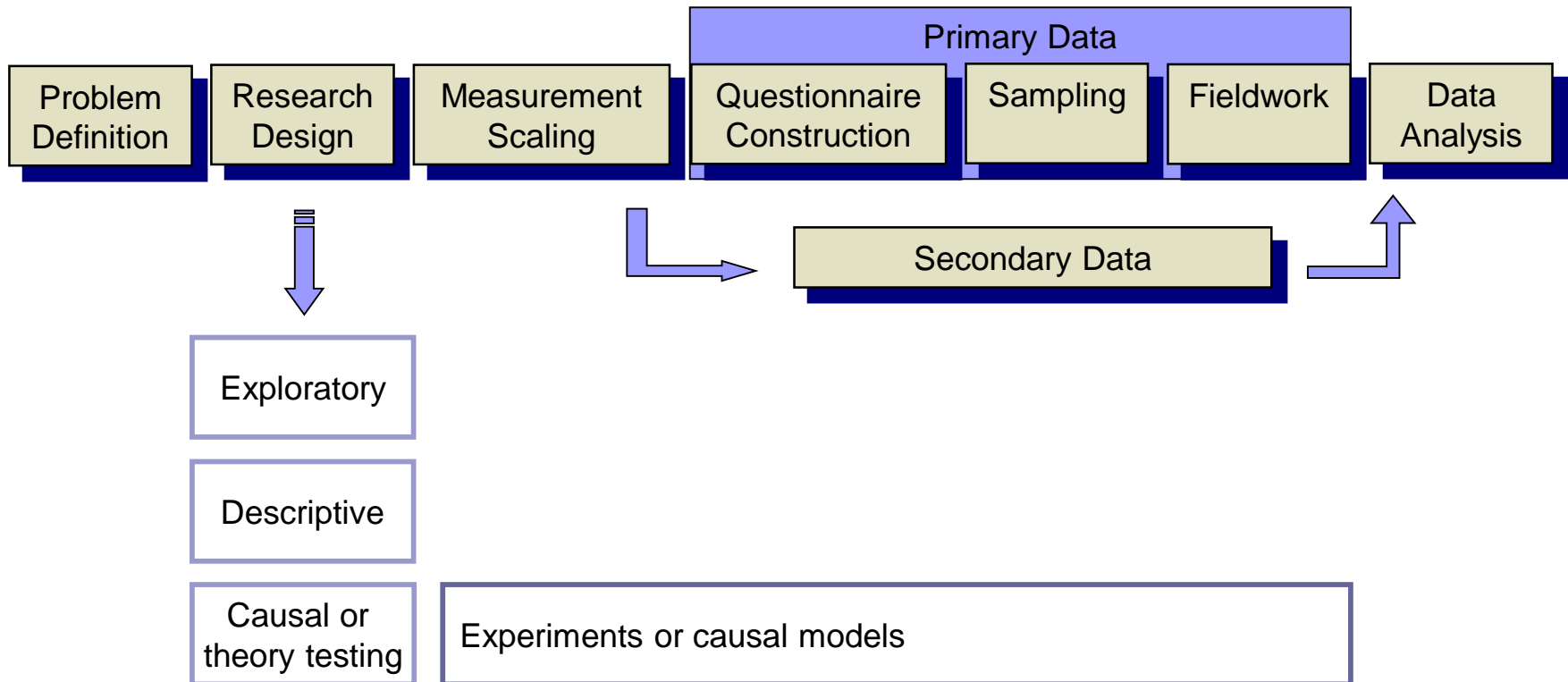
- Individuals in the distribution channels and trade associations
- Personal experience – speculation : first-step
- PROS: quick, first overview to guide future research
- CONS: highly subjective

# Marketing Research Process



- Direct observation of customers, their choices and behavior
- PROS: actual behavior, not reported
- CONS: collect data, research skills

# Marketing Research Process



- Scientific modeling techniques to understand a phenomenon
- Link decision variables (price, quality, ...) to an output variable (choice, quantity bought, RFM), traditionally from existing scanner DB, customer DB
- Mainly for promotion effectiveness, price sensitivity studies
- PROS: actual behavior, not reported
- CONS: collect data or manage DB, research skills, time-consuming

# Scanner Data in Practice

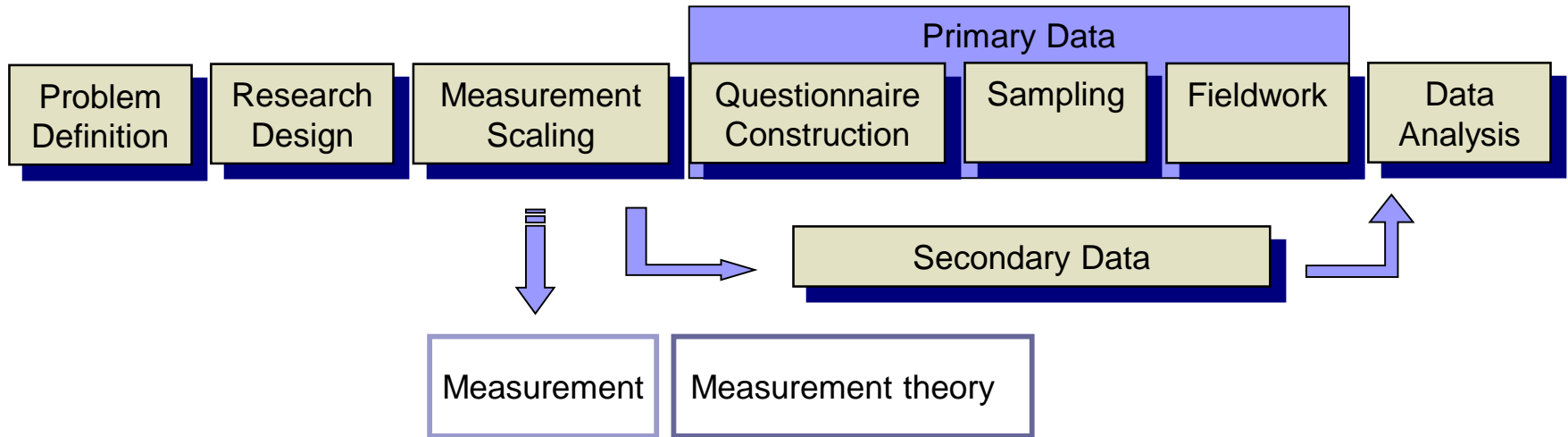
- *Practitioners reported that scanner data analysis has [...] been most widely adopted for decision making in consumer promotions (i.e., coupons), trade promotions, and pricing. For example, logit and regression models applied to scanner data have revealed very low average consumer response to coupons which has directly led to reduced couponing activity.*
- *In the case of new product decisions, scanner data use has been slow to develop due to the inherent limitations of historical data for these decisions.*
- *In advertising, scanner data is widely analyzed with models, but confusion among practitioners is very high due to controversies about methods and conflicting results.*

Bucklin and Gupta (1999), *Marketing Science*

# Use of Data Mining / Artificial Intelligence Techniques in Marketing

	Gini coefficient	Top decile
Binary logit model	0.241	1.775
Bagging	0.281	2.246
Stochastic gradient boosting	0.280	2.290

# Marketing Research Process





# Measurement Theory

Concept

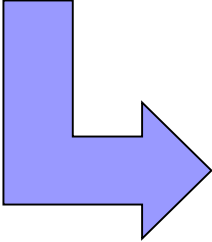
Service quality

Concept = An abstraction formed by generalization about particulars, e.g. mass, strength, love, advertising effectiveness, customer attitude.

Construct

Discrepancy between a customer's expectations for a service offering and the customer's perceptions of the service received

Construct = An observable, measurable concept (conscious inventions of researchers)

- 
- Service tangibles
  - Service reliability
  - Service responsiveness
  - Service assurance
  - Service empathy

# Measurement Theory

- Measurement:
  - The aim is to *measure*, i.e. assigning numbers to, a construct
  - The construct is measured using a set of observable indicators or *items* that form a scale, i.e. questions in a survey
- Scale of items: SERVQUAL scale (Parasuraman, 1988)
- We get a score for each of the items

# SERVQUAL: First Factor

## ■ Service Tangibles

- They should have up-to-date equipment.
- Their physical facilities should be visually appealing.
- Their employees should be well dressed and appear neat.
- The appearance of the physical facilities of these firms should be in keeping with the type of services provided.

Totally  
disagree

Totally  
agree

1

2

3

4

5

1

2

3

4

5

1

2

3

4

5

1

2

3

4

5

# SERVQUAL: Second Factor

## ■ Service Reliability

- When these firms promise to do something by a certain time, they should do so.
- When customers have problems, these firms should be sympathetic and reassuring.
- These firms should be dependable.
- They should provide their services at the time they promise to do so.
- They should keep their records accurately.

Totally  
disagree

Totally  
agree

1

2

3

4

5

1

2

3

4

5

1

2

3

4

5

1

2

3

4

5

1

2

3

4

5

# Measurement Theory

- Reliability = Constancy/consistency of the answers over time and across individuals
- Validity = Does the answers reflect the actual state / true answer

- $X_0$  = observed answer

- $X_T$  = true/actual answer

- $X_S$  = systematic error

- $X_R$  = random error

Perfect validity:

$$X_0 = X_T$$

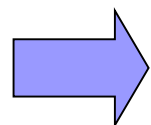
Not valid but

Perfect reliability:

$$X_0 = X_T + X_S$$

Not valid and not reliable:

$$X_0 = X_T + X_S + X_R$$



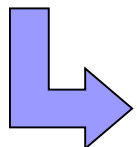
$$Var(X_0) = \underbrace{Var(X_T) + Var(X_S)}_{\text{Systematic variance}} + \underbrace{Var(X_R)}_{\text{Error variance}}$$

Systematic variance

Error variance

# Reliability

- Reliability coefficient (Constancy/consistency of the answers over time and across individuals)



$$r = \frac{\text{Var}(X_0) - \text{Var}(X_R)}{\text{Var}(X_0)} = 1 - \frac{\text{Var}(X_R)}{\text{Var}(X_0)}$$

The term  $\text{Var}(X_R)$  in the denominator of the second fraction is circled in blue. A blue arrow points from the circle to a question mark '?' to the right. Another blue arrow points from the circle back down and left towards the denominator of the first fraction.

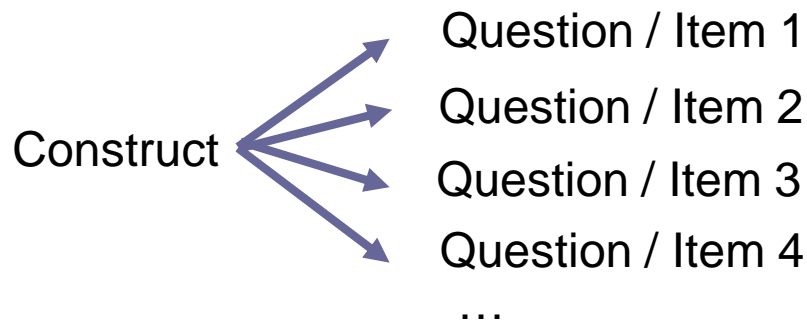
## Solutions?

- Ask the same questions several times → not feasible
- Ask different questions on the same “issue” or construct and measure the variance

→ internal consistency

# Reliability (II)

→ Internal consistency of the construct = Cronbach  $\alpha$



$$\alpha = \frac{k\rho}{1 + (k - 1)\rho}$$

$k$  = number questions/items

$\rho$  = average inter - item correlation

**! Cronbach  $\alpha$  is ideally higher than 0.7**

# Service Quality SERVQUAL

Dimension	Label	Number of Items	Reliability Coefficients (Alphas)	Items	Factor Loadings of Items on Dimensions to Which They Belong <sup>a</sup>
Tangibles	F1	4	.72	Q1	69
				Q2	68
				Q3	64
				Q4	51
Reliability	F2	5	.83	Q5	75
				Q6	63
				Q7	71
				Q8	75
				Q9	50
Responsiveness	F3	4	.82	Q10	51
				Q11	77
				Q12	66
				Q13	86
Assurance	F4	4	.81	Q14	38
				Q15	72
				Q16	80
				Q17	45
Empathy	F5	5	.86	Q18	78
				Q19	81
				Q20	59
				Q21	71
				Q22	68



# Cross-National Equivalence

- Is a construct's scale *equivalent* across countries?
  - The concepts, behaviors or objects should be interpreted equivalently: (*conceptual equivalence*)
    - E.g. concepts of “well-dressed”, “timely”, ...
  - Measures should be calibrated equivalently: (*calibration equivalence*)
    - Equivalence in monetary units, distance, volume, color, shape...
    - E.g. US \$ vs. £
  - Questions should be translated such that they remain equivalent: (*translation equivalence*)
    - Forward translation & back translation
  - Score should have the same interpretation: (*score equivalence*)
    - Equivalence of the observed scores on the measures
    - See Response biases on the next slide!

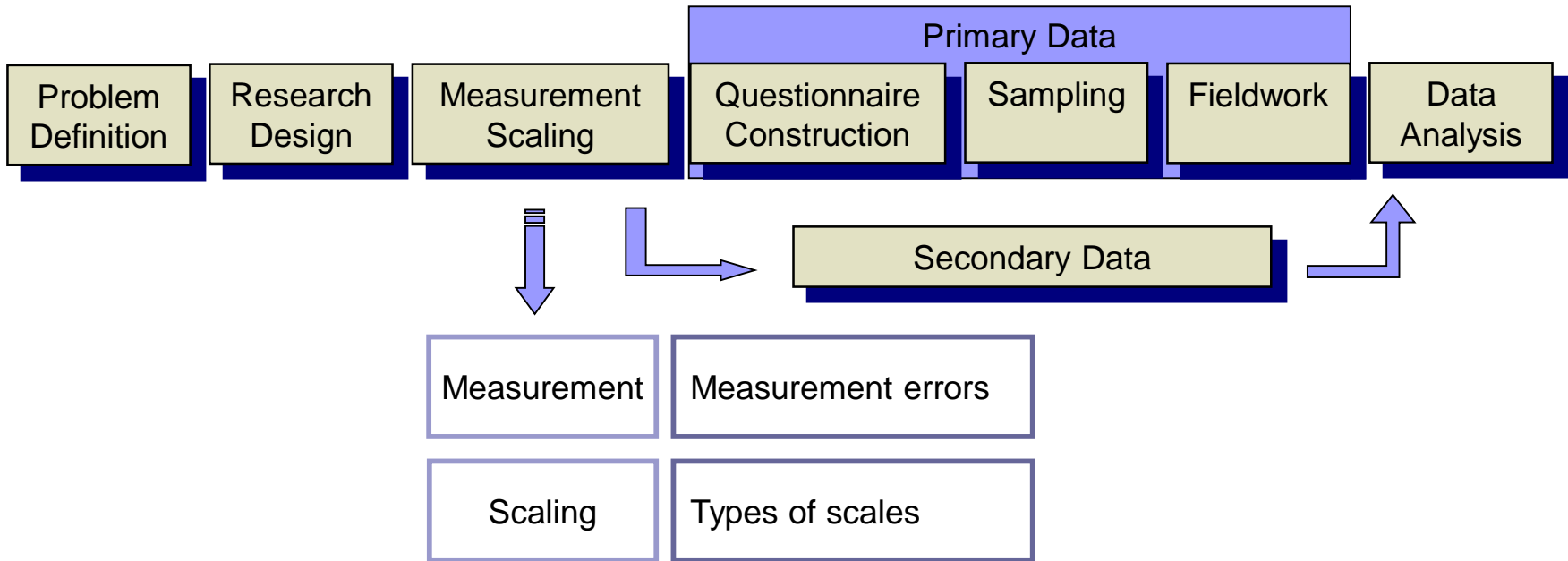
# Response Biases

- Respondents are influenced by content-irrelevant factors
- Response styles/biases:
  - **Social acquiescence or Yes-saying:** tendency to agree with items regardless of content
  - **Social disacquiescence:** tendency to disagree with items regardless of content
  - **Net acquiescence or directional bias:** tendency to agree more than disagree regardless of content
  - **Extreme responding:** tendency to endorse the most extreme response categories regardless of content
  - **Response range:** tendency to use a narrow or wide range of response categories around the mean
  - **Midpoint responding:** tendency to use the middle scale category regardless of content
  - **Noncontingent responding:** tendency to respond to items carelessly, randomly or non purposefully
  - **Social desirability:** tendency to give the socially desirable response rather than describe what they actually think, believe or do.
  - **Question order:** tendency to favor response categories because of their position in a list

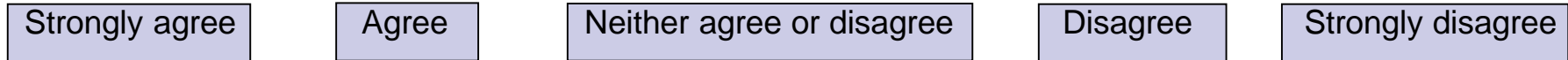
# Response Biases (II)

- Responses also depend on:
  - Respondent characteristics and personality traits
  - Self-referent cultural bias (interpretation by the researcher according to his/her own cultural referents)
  - Situational factors (e.g. time pressure)
  - Item non-response
  - Response format
  
- ➔ Solving response bias problems
  - Use positively worded and negatively worded items; balancing scales helps to control partially for acquiescence and disacquiescence

# Marketing Research Process



- Basic scales of measurement
  - Nominal: e.g. gender? Male or female
  - Ordinal: e.g. rank brand according to an attribute
  - Interval: e.g. visits frequency to that store
- One of the most well-known types of scale: Likert scales:



# Example

How do you feel about your family's financial condition over the past 6 months?

- Comfortable and secure
- Able to make ends meet
- Occasional difficulties
- Tough to make ends meet
- In over your head

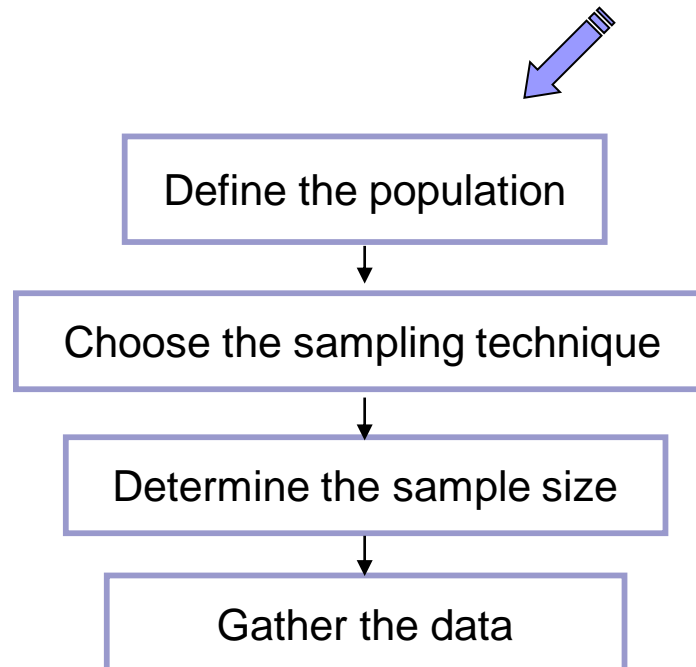
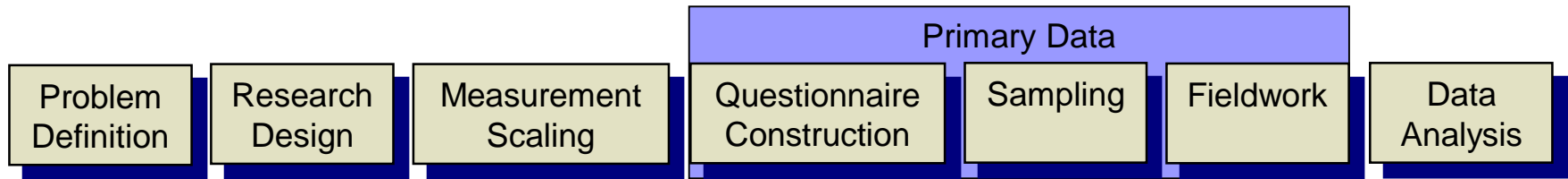
On a scale of 1 to 10, with 1 being not at all stressed and 10 being extremely stressed, how would you relate your financial stress level?

- 1  
Overwhelming Stress
- 2
- 3
- 4  
High Stress
- 5
- 6
- 7  
Low Stress
- 8
- 9
- 10  
No Stress at All

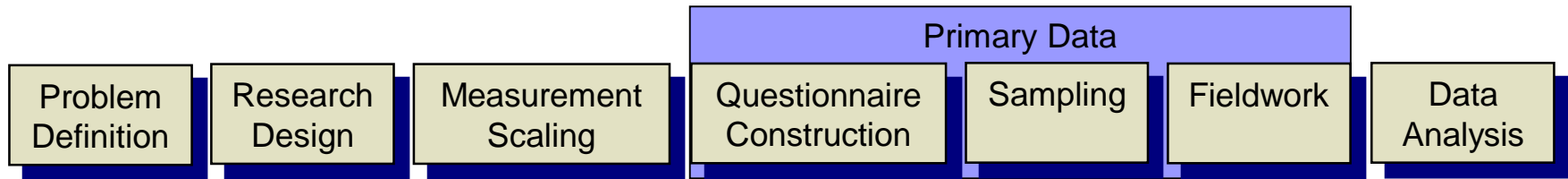
# Example (II)

How healthy are the following foods to you? Please share your feelings by rating them on the scale from <b>Very Unhealthy</b> ... to... <b>Very Healthy</b> .							
	Very Unhealthy	Moderately Unhealthy	Slightly Unhealthy	Not Healthy or Unhealthy	Slightly Healthy	Moderately Healthy	Very Healthy
Dark Chocolate	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Regular French Fries	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Baked potatoes from Potatoes high in Anti-oxidants	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Potato Chips (regular)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Broccoli	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
French Fries from Potatoes high in Anti-oxidants	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Baked Potatoes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
French Fries made from Fresh Potatoes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Potato Chips (baked)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

# Process of Sample Selection



# Process of Sample Selection



Choose the sampling technique

## 1. Probabilistic sampling

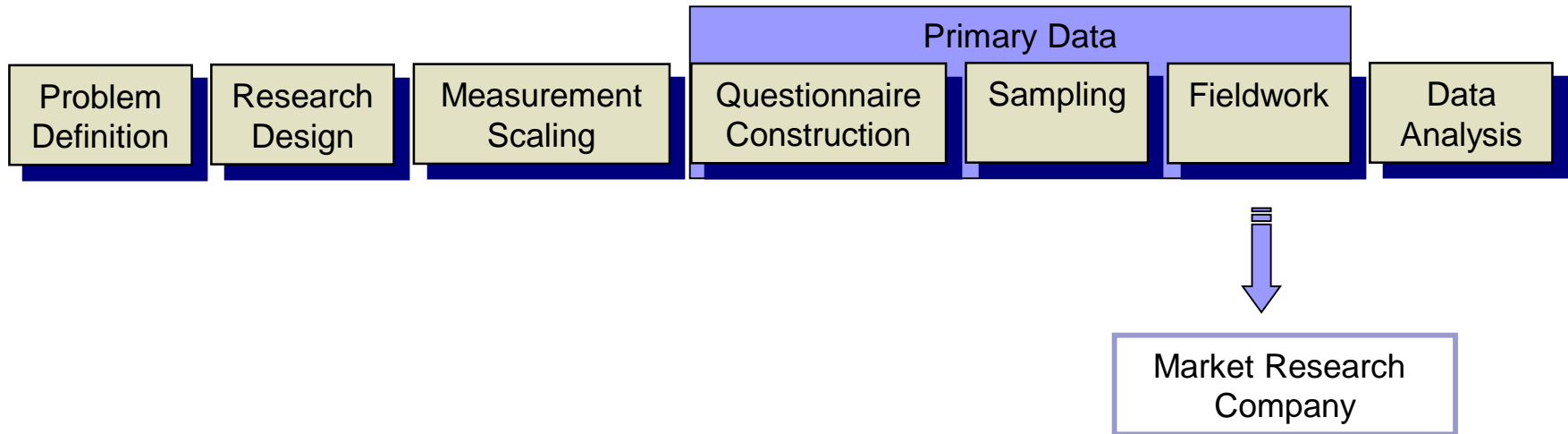
- Allows for bias-free selection of sample units
- Permits the measurement of sampling error

## 2. Non-probabilistic sampling

- Relies on the expertise of the person taking the sample
- Convenience: the researcher chooses the respondents on a convenient place (typical cities chosen as test markets, “intercept” respondents in shopping malls)
- Judgment: the researcher chooses the respondents based on some a priori criteria
- Quota: the researcher establishes some quota for the various types of persons he wants to include in the sample (e.g. Bureau of the Census data)
- Snowball: the researcher asks respondents to provide the names of additional respondents to be included.



# Fieldwork

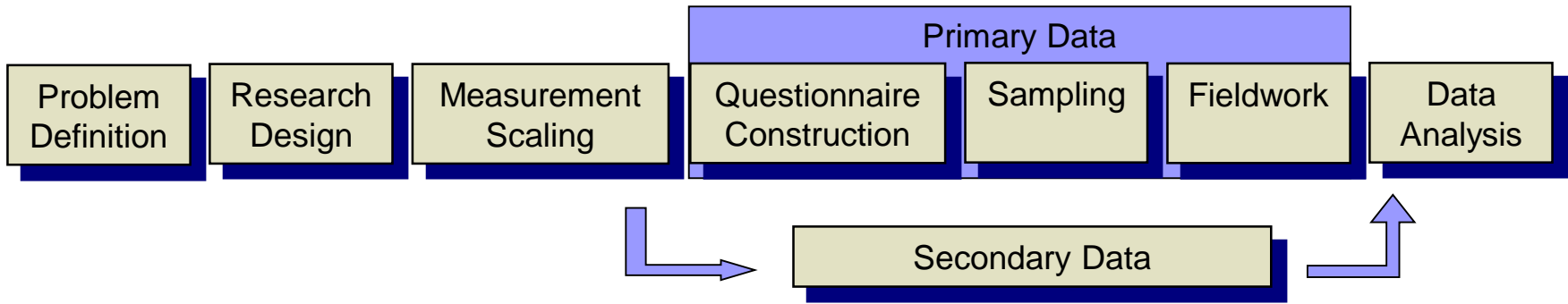


## Top Global Research Companies

1. VNU – AC Nielsen (NL) [www.vnu.com](http://www.vnu.com)
2. Taylor Nelson Sofres (UK) [www.tns-global.com](http://www.tns-global.com)
3. IMS Health (US) [www.imshealth.com](http://www.imshealth.com)
4. Kantar Group (US) [www.kantargroup.com](http://www.kantargroup.com)
5. Gfk Group (GE) [www.gfk.com](http://www.gfk.com)
6. Ipsos Group (FR) [www.ipsos.com](http://www.ipsos.com)

...

# Secondary Data



## Main Secondary Data Sources

See [www.internetworldstats.com/sources.htm](http://www.internetworldstats.com/sources.htm)

US Department of Commerce

[www.stat-usa.gov](http://www.stat-usa.gov)

GlobalEDGE – Michigan State University

[www.ciber.msu.edu](http://www.ciber.msu.edu)

United Nations [www.un.org](http://www.un.org)

World Bank [www.worldbank.org](http://www.worldbank.org)

OECD [www.oecd.org](http://www.oecd.org)

EU Commission [www.europa.eu.int](http://www.europa.eu.int)

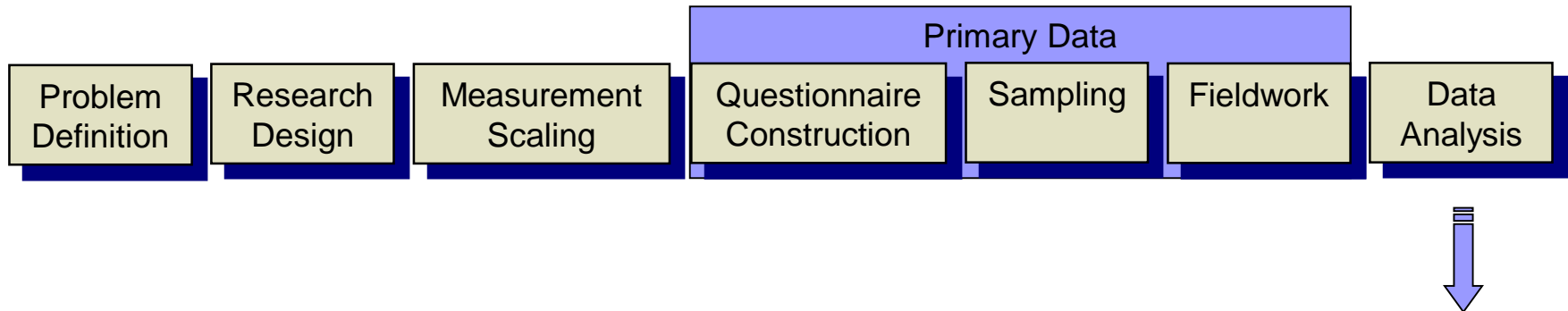
CIA World Factbook

IMF International Financial Stats

Euromonitor [www.euromonitor.com](http://www.euromonitor.com)

- Ready availability
- Low cost compared to primary data
- As a first step (overview) to be completed by detailed primary data
- Data accuracy and equivalence issues

# Data Analysis



- Cross-tabulation
- T-tests
- Analysis of variance
- Analysis of covariance
- Multivariate analysis of covariance
- Multiple regression
- Hierarchical linear models
- Multiple discriminant analysis
- Correlation analysis
- Cluster analysis
- Multidimensional scaling
- Factor analysis
- Confirmatory factor analysis
- ...

See *“Market Research and Analysis”* (Dr. Vijay)  
All methods are available in SPSS