Columbia Business School

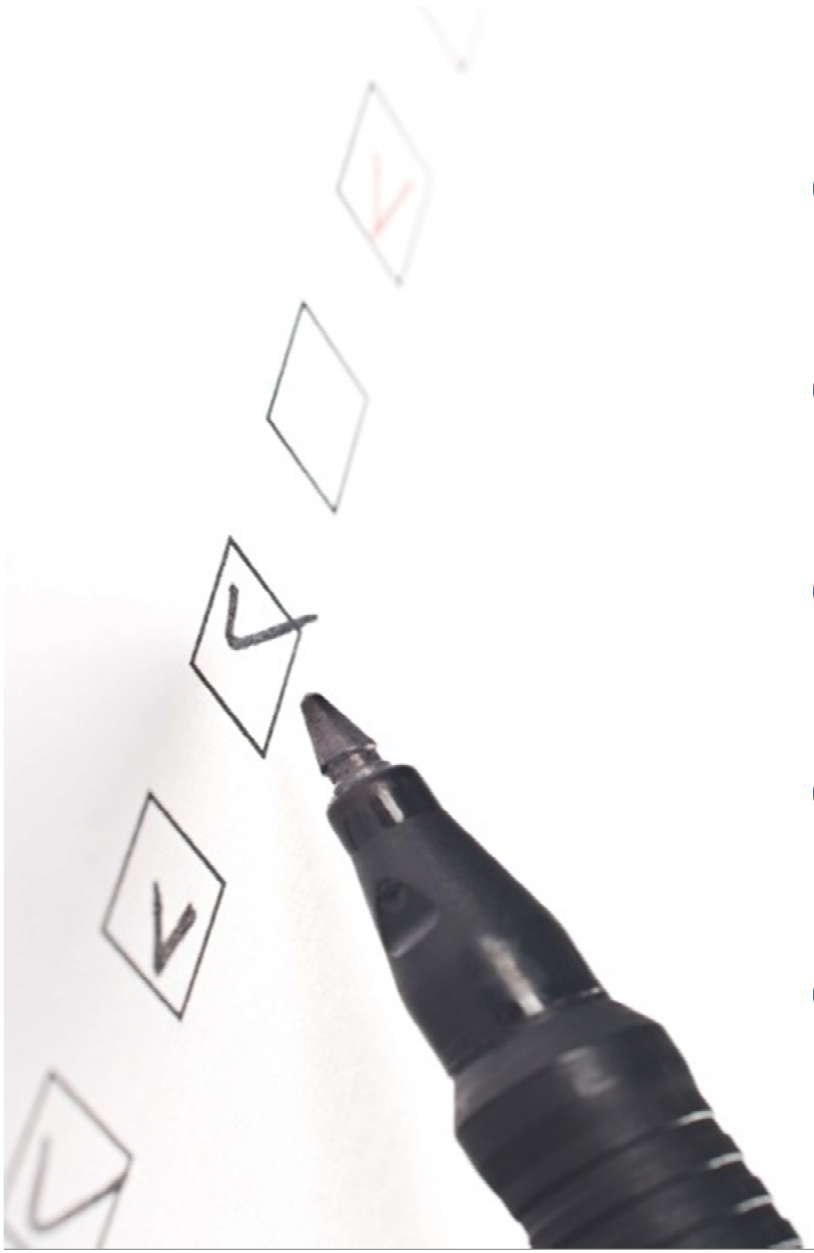
Retail & Co. – Driving Retail Excellence

Applied Regression Analysis – Final Project

Group 3: Buaron, Costa, Gliklich, Pozzi, Ricci, Tribioli

July, 1st 2013

Objectives for today's presentation



1. Describe the reason for our engagement
2. Define the situation and the approach we followed
3. Present the results of our study
4. Compare management vs. customer view
5. Share further developments

The client is a haute-couture retailer with its core business in fashion and leather goods

COMPANY DESCRIPTION

STRONG HERITAGE

*"We were **founded in the mid-40s in Paris**. Since then we've always been on the **top of the fashion/ luxury scene**."*

Retail Director, Client

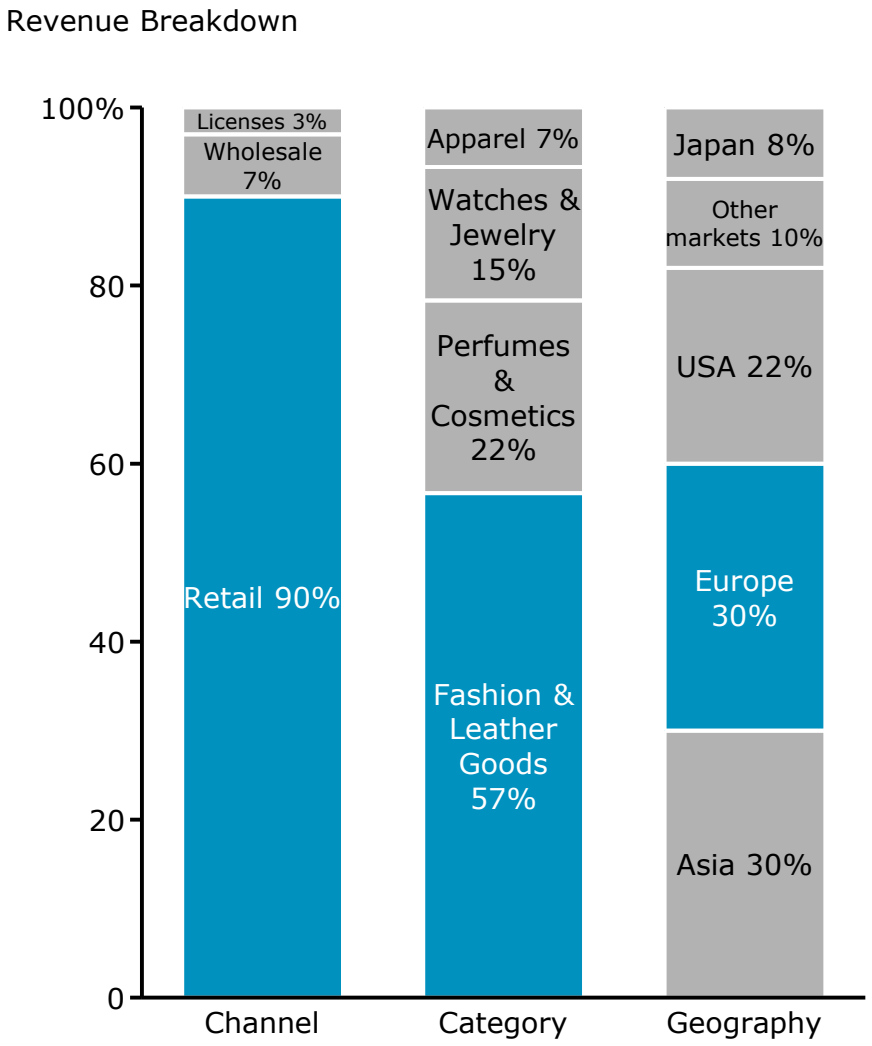
GENERAL INFO

- **Revenues 2012**: more than €500M
- **EBITDA Margin 2012**: ~20%
- **Retail locations**: 210 shops all over the world
- **Average price tag**: ~€900

ICONIC PRODUCTS



BUSINESS MODEL DESCRIPTION



Source: official filings, 2012; web literature

The goal of our project: reaching operational excellence through an advanced management tool

WHAT WE HEARD FROM THE COMPANY…

TWO KEY STAKEHOLDERS FROM THE CLIENT SIDE

- The **Southern Europe Retail Director**, worried about operational excellence
- The **Business Development Director**, specifically concerned about retail results forecasts

…WHAT WE ARE SUPPORTING THEM WITH

DEVELOP A MANAGEMENT TOOL TO DRIVE IMPROVED RETAIL OPERATIONS

- **Identify** the **best predictors for sales per square meter**
- Provide the management with **decision tools to enhance operational excellence**
- **Validate** the **current store segmentation**

Source: management interviews, June 2013

We considered variables about stores and their sales forces as a means to predict Sales per Square Meter

VARIABLES DESCRIPTION

- The key metric by which stores are evaluated in SALES PER SQUARE METER which we took as our dependent variable.
  - -Across the 61 stores we considered, sales per square meter ranged from €7,700 to €22,000 with an average of €14,500
- The independent variables we considered could be broadly bucketed into two buckets - Store Characteristics and Sales Force – but which are highly interrelated

STORE CHARACTERISTICS		SALES FORCE
CONCEPT	SIZE	NUMBER OF SALES ASSISTANTS
YEARS SINCE OPENING/ RENOVATION	CARRYING ENTIRE COLLECTION	SALES ASSISTANTS PER SQUARE METER
NUMBER OF ROOMS	NUMBER OF WINDOWS	PERCENTAGE OF CUSTOMERS ENGAGED

We can anticipate some 'diseases' in our data from a correlation matrix

CORRELATION MATRIX

	DEPENDENT VARIABLE Sales per Square Meter	Concept: Corner	Concept: Standard	Concept: Flagship	Years Since Opening / Refurbishment	Size	Square Meters per Assistant	Number Sales Assistants	Number of Rooms	Customer Engagement	Carrying Entire Collection	Number of Windows
Sales per Square Meter	1.00											
Corner	0.84	1.00										
Standard	(0.19)	(0.61)	1.00									
Flagship	(0.66)	(0.33)	(0.55)	1.00								
Years Since Opening / Refurb	(0.61)	(0.48)	0.15	0.33	1.00							
Size	(0.95)	(0.78)	0.04	0.77	0.50	1.00						
Assistants per Square Meter	(0.58)	(0.78)	0.62	0.08	0.30	0.49	1.00					
Number of Sales Assistants	(0.86)	(0.68)	(0.15)	0.88	0.44	0.91	0.35	1.00				
Number of Rooms	(0.86)	(0.70)	(0.06)	0.81	0.48	0.89	0.41	0.97	1.00			
Customer Engagement	0.96	0.84	(0.15)	(0.70)	(0.57)	(0.94)	(0.55)	(0.89)	(0.89)	1.00		
Carrying Entire Collection	(0.62)	(0.59)	0.05	0.55	0.35	0.66	0.37	0.67	0.66	(0.63)	1.00	
Number of Windows	(0.87)	(0.72)	(0.08)	0.85	0.44	0.91	0.42	0.99	0.95	(0.90)	0.68	1.00

CUSTOMER ENGAGEMENT CREATES MULTICOLLINEARITY ISSUES WITH MANY OF OUR VARIABLES

The Best Subset regression explains 96.1% of Sales/Sqm. However, there is multicollinearity

VARIABLES

- CONCEPT
- CARRYING ENTIRE COLLECTION
- SIZE
- NUMBER OF WINDOWS
- YEARS FORM RENOVATION
- NUMBER OF ROOMS
- PERCENTAGE OF CUSTOMERS ENGAGED
- SALES ASSISTANTS PER SQUARE METERS
- NUMBER OF SALES ASSISTANTS

BEST SUBSET ANALYSIS

RESULTS

INPUT VARIABLES

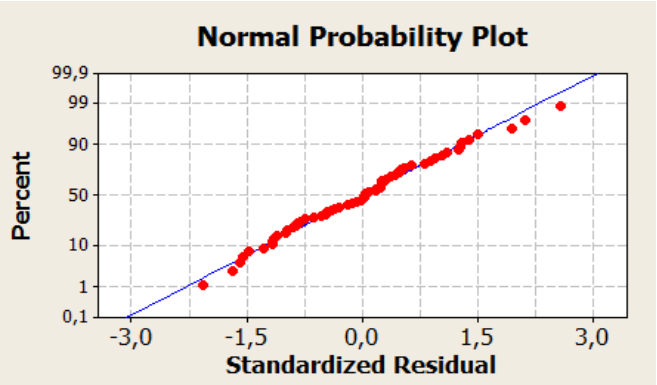
- CONCEPT
- SIZE
- YEARS FROM RENOVATION
- ASSISTANT/SQM
- PERCENTAGE OF CUSTOMERS ENGAGED
- NUMBER OF SALES ASSISTANTS

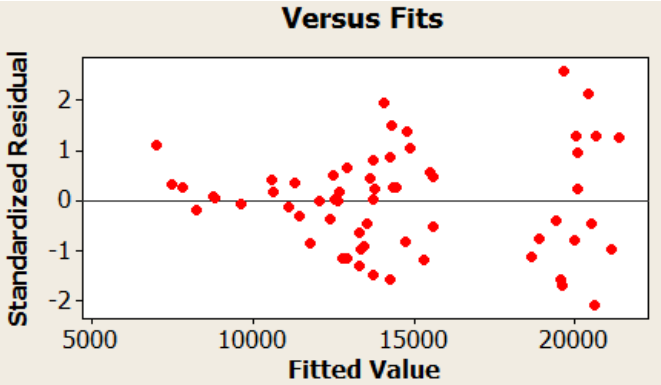
REGRESSION RESULTS

R^2 = 96.1%
R^2 Adjusted = 95.6%

	Coefficients	*P-value*	*VIF*
Intercept	9919	0,00	
Years from renovation	-528	0,00	1,53
Selling Sqm	-24	0,00	13,07
Sales Assistants	-136	0,26	20,34
Customer Engagement	12994	0,00	15,2
Corner	-1105	0,27	16,56
Flagship	2127	0,04	16,03
Assistant/sqm	66720	0,18	3,33

REGRESSION PLOTS





The Best Subset regression explains 96.1% of Sales/Sqm. However, there is multicollinearity

HETERO-SCEDA-STICITY!!

Customer engagement and number of sales assistants are main sources of colinearity

DEPENDENT VARIABLE	BEST PREDICTORS	WEIGHT
CUSTOMER ENGAGEMENT	• Sales assistant • Size • Shopping areas • Concept • Other variables not available -Assistants churn -Customer traffic	
SALES ASSISTANTS	• Size of the store • Assistants per square meter	
STORE CONCEPT	• Size of the store	

Customer Engagement itself predicts a great portion of the variability in Sales per Square Meter

REGRESSION OF SALES/SQM VERSUS CUSTOMER ENGAGEMENT

R^2 = 92.7%
R^2 Adjusted = 92.6%

	Coefficients	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-171	0,76	-1279	937
Customer	24859	0,00	23045	26673

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regressio	1	862.911.055	862.911.055	752	0,000000
Residual	59	67.713.304	1.147.683		
Total	60	930.624.360			

AGE, SELLING SQUARE METERS AND ASSISTANTS/SQUARE METERS GIVE SIMILAR RESULTS

R^2 = 93.6%
R^2 Adjusted = 93.3%

	Coefficien	*P-value*	*VIF*
Intercept	16740	0,000	
Years from renovation	-807	0,000	1,34
Selling Sqm	-36	0,000	1,68
Assistant/sqm	134281	0,001	1,4

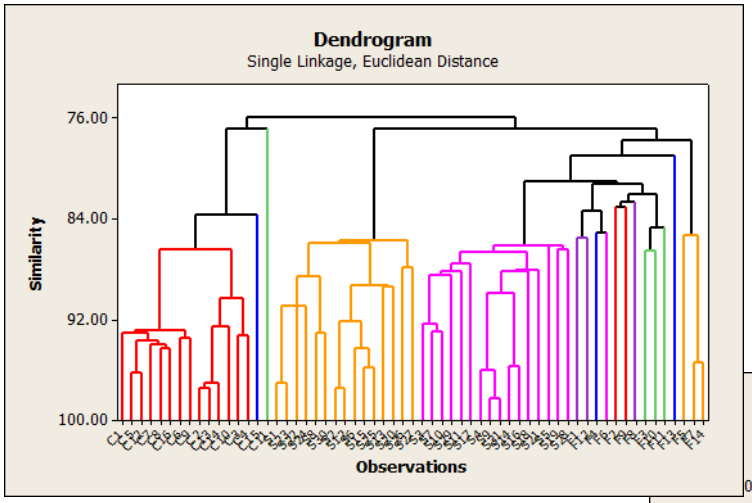
ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	871.091.197	290.363.732	278,01	0,000000
Residual	57	59.533.163	1.044.441		
Total	60	930.624.360			

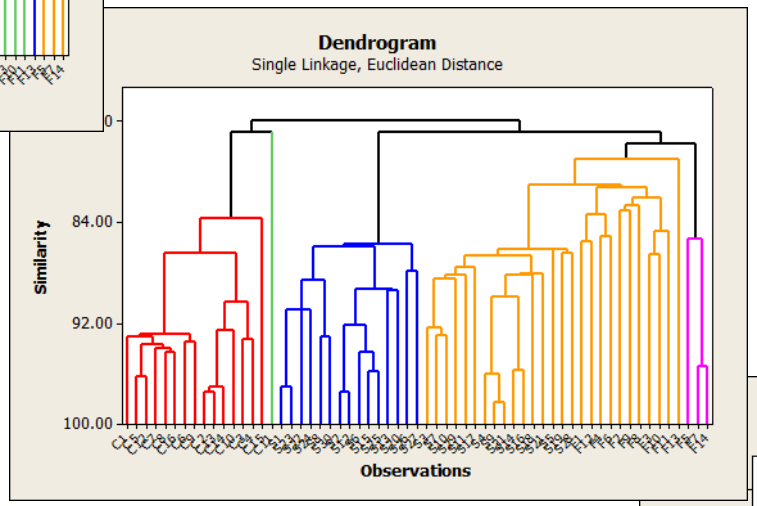
BEST FIT MODEL

...SIMILAR RESULTS CAN BE ACHIEVED COMBINING OTHER VARIABLES

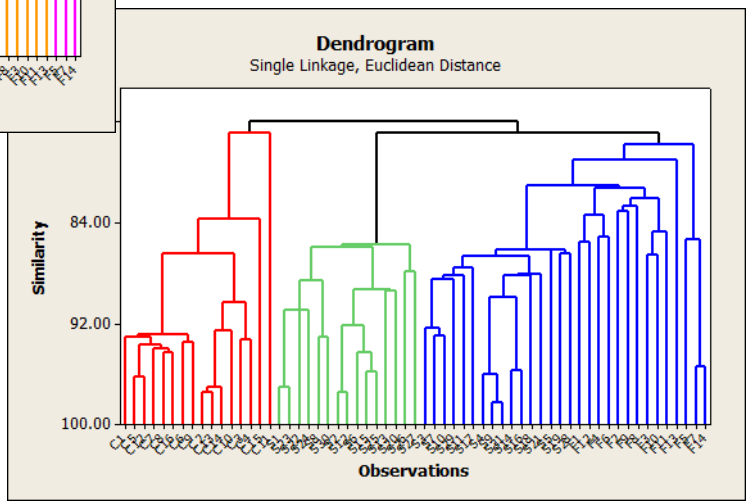
A cluster analysis helped us understand the importance of store formats



REDUCED # OF CLUSTERS FROM 15 TO 5



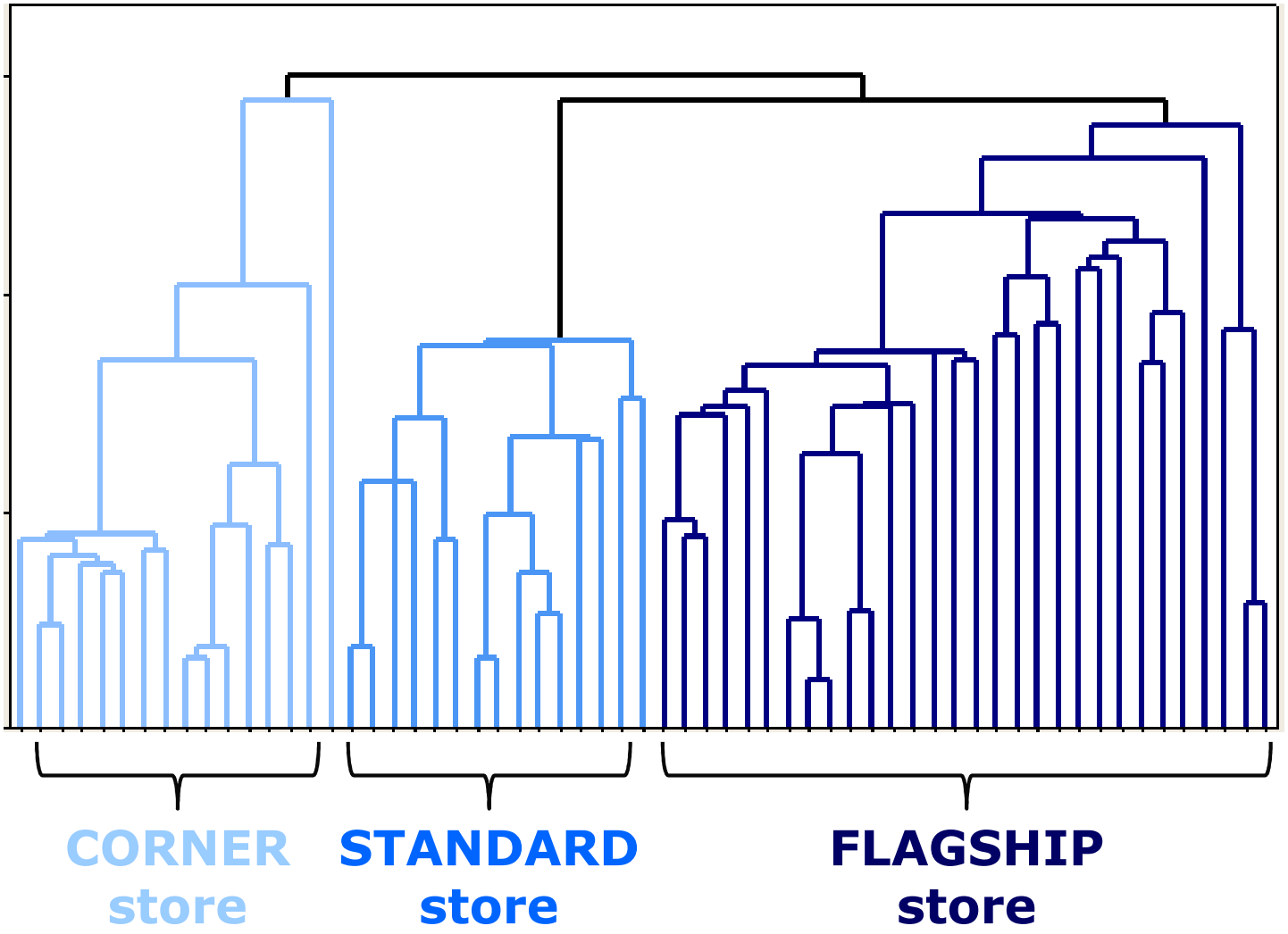
REDUCED # OF CLUSTERS FROM 5 TO 3



*"After analyzing the performance of each concept, we **came out with what we believed was the most successful store formats** and **retail results proved us right**."*

Business Development Director, Client

The 3-clusters analysis split up our stores into the three different store formats



IF THE SOLVER IS FORCED TO IDENTIFY ONLY 2 CLUSTERS, STANDARD AND FLAGSHIP STORES ARE GROUPED TOGETHER

Store concepts are a perfect fit for the 3 clusters among our 61 stores

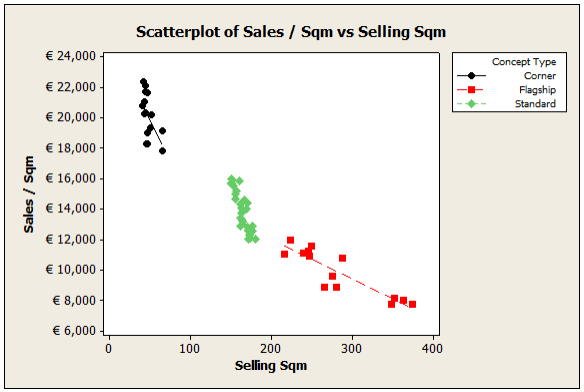
	SALES/ SQM	SELLING SQM	# OF WINDOWS	CUSTOMER ENGAGEMENT
CORNER Store	20’054€	48 sqm	0.3	81%
STANDARD Store	13’824€	165 sqm	2.7	57%
FLAGSHIP Store	9’820€	284 sqm	6.3	40%

TO IMPROVE SALES/SQM ON STANDARD AND FLAGSHIP STORES OUR CLIENT CAN FOCUS ON IMPROVING CUSTOMER ENGAGEMENT

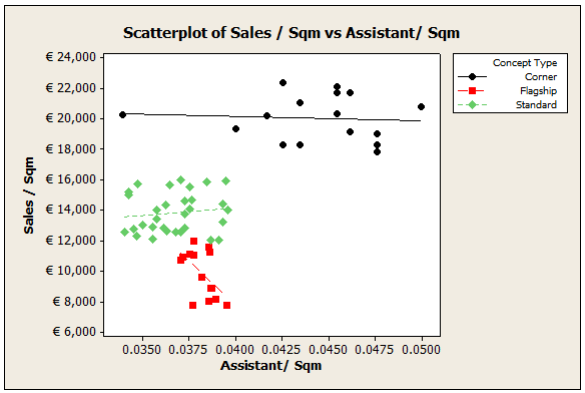
Interaction variables do not significantly help to improve the model

Interaction variables analysis

- We want to see if it is possible to further improve the model, by using **interaction variables**
- There seem to be some "interaction" between **Sales/sqm** and
  - **Selling sqm**



  - **# assistant/sqm**



Outcome

- We decided **not to include the interaction variables** because they

⊕ **Increase adjusted R^2** by only by 0.3%

⊖ **Increase complexity** of the model, therefore is

⊖ Do **not** significantly **improve** the distribution of **residuals**

⊖ **Multicollinearity issues** between the interaction variables and the explanatory variables already included in the model

We advise management to use the most *actionable* model

Recommended model

	Model 1 **REGRESSION OF SALES/SQM VERSUS CUSTOMER ENGAGEMENT**	Model 2 **AGE, SELLING SQUARE METERS AND ASSISTANTS/SQUARE METERS GIVE SIMILAR RESULTS**
STATISTICAL CONSIDERATIONS	Both models are **good and robust**: - Very low p-value of the F-test - High explanatory power (R^2) - Very low p-value of the t-test - Low VIF factors	
	⊖ **Residual highly skewed** and not homogenously distributed	⊕ Residuals **follow** more closely a **normal distribution**, even though not perfectly. We run a good**ness of fit test**
	⊕ Very **simple** and **straightforward** model	⊖ **Slightly more complex** model
MANAGERIAL CONSIDERATIONS	⊖ **Not actionable**, difficult to understand which actions should be taken	⊕ **Actionable**, management can take actions based on the model
	⊖ **Less useful** to make **prediction**, customer engagement is not objective and easy to forecast as variables in model 2	⊕ Easier to use to make **prediction**, since all variables used in the model are objective and easily predictable

Both the regression model and the customer survey identify assistants as the most important variable

IN-STORE SURVEY RESULTS

Source: client in-store interviews, n=698, Jan-Apr 2013

CUSTOMER PERSPECTIVE

"***Without my preferred retail-associate I would be lost***. *She is my* ***point of reference*** *in terms of style and new trends."*

In-store customer 1

"***I trust the advices of the retail-associates***. ***Year after year***, *they have been taking care of my wardrobe and I have always been classy and stylish."*

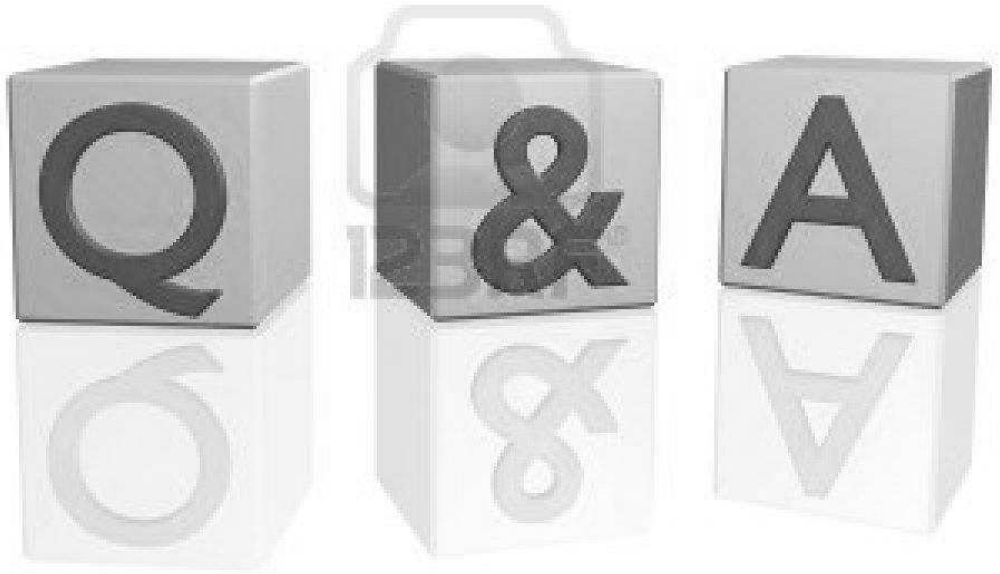
In-store customer 2

CUSTOMER ENGAGEMENT IS ALSO THE MOST IMPORTANT FACTOR FROM THE CUSTOMER PERSPECTIVE

Next steps: how can we take this further to support even better the management?

POTENTIAL FURTHER DEVELOPMENTS	INFORMATION/DATA REQUIRED
• Expand the geographic focus on all countries of retail presence	• Gather retail store data from other geographic regions
• Develop a management tool for each of the store formats	• Collect additional retail store data per concept
• Complement the management tool with a prediction model for future store openings	• Collect demographic, traffic competition, and location information

Q&A: we are here to answer your questions… ☺

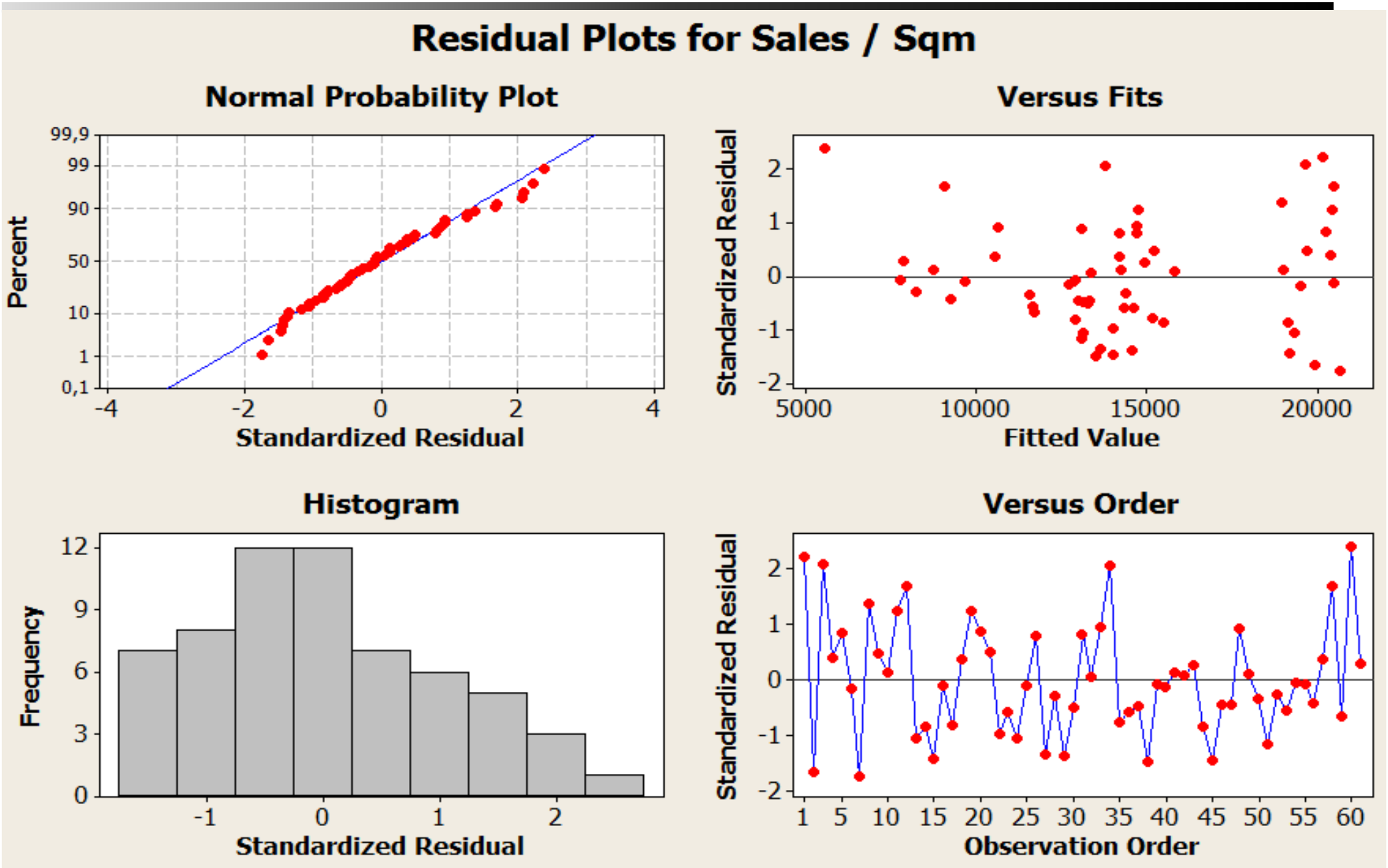


APPENDIX

Overview of variables and variability

	Mean	Standard Deviation	Range
Sales Per Square Meter	$14,539	$3,938	$7,742 - $22,365
Year Opened or Refurbished	2012	0.87	2010 - 2013
Age	2.13	0.87	1 - 4
Size (in Square Meters)	161	87	40 - 374
Square Meters per Assistant	26	2	20 - 30
Number of Sales Assistant	6.64	4.02	2 - 15
Numer of Rooms	2.30	1.12	1 - 5
Customer Engagement	59.2%	15.3%	31.5% - 88.7%
Number of Windows	2.89	2.21	0 - 7

Residual plots
Sales/sqm vs. Age, Assistants/Sqm and Selling Sqm



Variables average values differ among customers

	SALES ASS/ SQM	YEARS FROM RENOVATION	SHOPPING AREAS	WINDOWS	ALL COLLECTION
CORNER Store	0.045	1.4	1.0	0.3	-
STANDARD Store	0.037	2.3	2.2	2.7	0.5
FLAGSHIP Store	0.048	2.6	3.9	6.3	1