Rules and restrictions for publications

"Papers on normalization, variable selection, classification or clustering of microarray data"

Examples of restrictions by Bioinformatics for

New methods for these problems to meet a very high standard, showing important improvement in results for real biological data, as well as novelty

News to consider when publishing : Rules and Restrictions

Guidelines and advices on

- Preparing and giving a presentation
- Preparing a written report

Exercises on Paper 1 to be discussed

Examples of Bad presentation (by A.Regl)

1. Gates versus Jobs
2. Microsoft corporative presentation

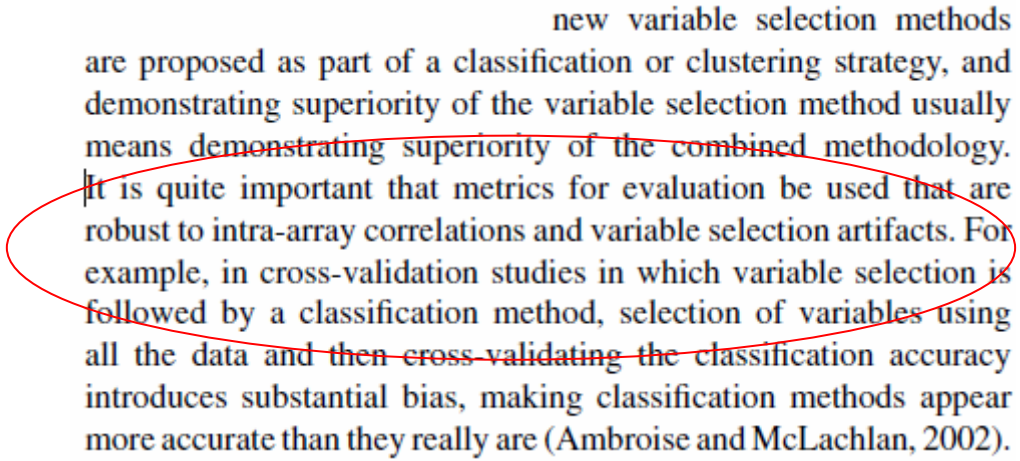
Rules and restrictions for publications

Standards before submission of papers to *Bioinformatics*
The role of simulation
Normalization

Normalization: Normalization necessarily involves a trade-off between its positive role in reducing variability, and its potentially negative role in increasing bias. There are a number of good image analysis, preprocessing, transformation and normalization methods extant for single- and dual-color DNA microarrays. To show that a new method is better requires comparison demonstrating that results in differential expression analysis, classification or clustering are better with the new normalization method than with previous methods. Not one but several previous methods should be chosen for comparison including the most widely used approaches. Several datasets should be used, including spike-in and dilution studies when feasible, as well as 'real' biological datasets.

Rules and restrictions for publications

Variable selection



new variable selection methods are proposed as part of a classification or clustering strategy, and demonstrating superiority of the variable selection method usually means demonstrating superiority of the combined methodology. It is quite important that metrics for evaluation be used that are robust to intra-array correlations and variable selection artifacts. For example, in cross-validation studies in which variable selection is followed by a classification method, selection of variables using all the data and then cross-validating the classification accuracy introduces substantial bias, making classification methods appear more accurate than they really are (Ambroise and McLachlan, 2002).

Classification and prediction

Substantial improvement on the accuracy

Variable selection and parameter choice for all methods needs to be done strictly in the training set

Rules and restrictions for publications

Classification and prediction

Experience shows that there is considerable noise in classification accuracy experiments, so modest increases in achieved accuracy are usually not convincing. Experience also shows that classification performance in a microarray problem depends strongly on the dataset, and less on the variable selection and classification methods. More than modest differences are required to excite interest in a new method. Authors should keep in mind the 'No Free Lunch Theorems' of Wolpert and Macready (1997) which demonstrated that there is no optimization/classification method that outperforms all others in all circumstances (Wolpert, 1996).

Clustering

No ground truth against which to compare

Rules and restrictions for publications

Clustering

No ground truth against which to compare

> Ideally, a new clustering method would demonstrate novel biological insights or some attractive statistical properties not available from previous methods, including several commonly used methods. Requiring new biological findings is a difficult standard, but a necessary one to insure that new published methods are useful and likely to be used.

Communication and Presentation
Preparing and giving a talk

Saskia Litière & Monique Maelstaf
Center for Statistics Hasselt University

Overview

- ✓ Preparing a presentation
- ✓ Structure of a presentation
- ✓ Time management
- ✓ Presentation style
- ✓ Examples of bad slides
- ✓ Visual aids and other figures
- ✓ Giving and receiving criticism
- ✓ Conclusion

Preparation

- ✓ Goal of the presentation
  - Results of data collection
  - Action to be taken
  - Focus on goal
- ✓ Background of audience
- ✓ Allocated time
- ✓ Facilities

Preparation

Knowledge of content

The information has to be accurate, interesting, relevant and includes facts to support main ideas

Listeners have one chance to hear the presentation and cannot re-read when they get confused

Presenter should be instructive and mind-provoking

Preparation

Creativity

In a presentation, the audience wants to know the importance of what is being presented. Therefore, a good oral presentation should include applications in our day to day lives. This is because people tend to be more attentive to things which can help improve their doubts.

An important criterion is to understand the audience that one will present to. Are they scientists, engineers, business men/women, and so on. This will give insight into the nature of language to be used in the presentation

Structure

- ✓ Introduction
- ✓ Content table
- ✓ Methodology
- ✓ Results
- ✓ Conclusions
- ✓ Further research
- ✓ Questions

Structure and time management

- ✓ Tell the audience when you will allow questions
- ✓ Make sure you understand the question before answering, thus if you're not sure ask for some clarification
- ✓ Don't rush the question; answer it once the person who is asking is finished
- ✓ Be precise
- ✓ Avoid running over the given time. Plan the time distribution for different section
- ✓ A lengthy presentation makes the audience get bored, tired, and consequently unable to understand the rest of the work.

Eye Contact

Make frequent eye contact with the audience. Really look at the audience as you talk to them. Engaging them directly with your eyes transfers a bit of your energy to them and keeps them focused on your content. Making eye contact says that you are in charge of the room and for a presentation that's what you want.

Be enthusiastic

Show vibrant interest in the topic and eagerness to communicate. Do not stand in a 'frozen' position; move about but not too excessively

Cultural conventions

Use appropriate greetings and mode of dressing or appearance

Style

- ✓ Be relaxed and enthusiastic
- ✓ Avoid clutter
- ✓ Make eye contact with audience
- ✓ Speak loud enough
- ✓ Speak at a comfortable rate
- ✓ Avoid tics …..

Examples of bad slides ...

README.TXT

Do not attempt to put all the text, code, or explanation of what you are talking about directly onto the slide, especially if it consists of full, long sentences. Or paragraphs. There's no place for paragraphs on slides. If you have complete sentences, you can probably take something out.
If you do that, you will have too much stuff to read on the slide, which isn't always a good thing.
Like the previous slide, people do not really read all the stuff on the slides.
That's why it's called a "presentation" and not "a reading" of your work
Practice makes perfect, which is what gets you away from having to have all of you "notes" in textual form on the screen in front of you.
Utilize the Notes function of PowerPoint, have them printed out for your reference.
The audience doesn't need to hear the exact same thing that you are reading to them.
The bullet points are simply talking points and should attempt to summarize the big ideas that you are trying to convey
If you've reached anything less than 18 point font, for God's sake, please:
Remove some of the text
Split up the text and put it on separate slides
Perhaps you are trying to do much in this one slide?
Reading a slide is annoying.
You should not simply be a text-to-speech converter.

Font Size

You are close to your monitor

Your audience is far from the screen

Tahoma	TNR	Courier	Comic	Lucida Sans
32 pt	32 pt	32 pt	32 pt	32 pt
28 pt	28 pt	28 pt	28 pt	28 pt
24 pt	24 pt	24 pt	24 pt	24 pt
20 pt	20 pt	20 pt	20 pt	20 pt
18 pt	18 pt	18 pt	18 pt	18 pt
16 pt	16 pt	16 pt	16 pt	16 pt
14 pt	14 pt	14 pt	14 pt	14 pt
12 pt	12 pt	12 pt	12 pt	12 pt
10 pt	10 pt	10 pt	10 pt	10 pt

This is a really long title for this single slide, I should have just summarized the title in order to make it clear for the audience

Hard to read

Many people don't read the title anyway

Could have been "Long Slide Titles"

Know Slide Boundaries

People can't read text that runs off the side of the slide like the one I am typing right now in her

Bullets Aren't Everything

How many
Levels of
Hierarchy do
You think
You need
* To express
- Your point?

Pixar Studios

Previous slide(s) used “animation”…

Can

Use it sparingly

(it can be annoying)

Use only where it is USEFUL

Know if presentation system will handle

- Different versions of PowerPoint, Macs, etc.

Or use multiple slides to safely animate

- Flip-book style

Be Very

Animation

Distracting

Another animated movie

1958 world exhibition
Brussels
105 meters high
Iron atom 165 billion times enlarged …

The Atomium

Can you look at this for 45 minutes?

Colors look different on every LCD projector

Colors look different between transparencies and projector

Side note: if printing slides, may want to choose white background to save ink!

Mommy, my eyes are burning!

I See A Ghost

More contrast on monitor than projector
Different projectors == different results
Colors to avoid with white are:

- Light Green
- Light Blue
- Pale Yellow

Your slides should have good contrast

Usually can't read this…

Equations

$$
\begin{aligned}
X' &= A * B \\
&= (A - (2^p - 1)) * (B - (2^q - 1)) \\
&= AB - B(2^p - 1) - A(2^q - 1) + (2^p - 1)(2^q - 1) \\
X'' &= (A - E_p)(B + Eq) \\
&= AB + AE_q - BE_p - E_pE_q \\
&= AB + AE_q - (BE_p + E_pE_q) \\
&= AB + AE_q - \frac{E_pE_q}{2} - \left(BE_p + \frac{E_pE_q}{2}\right)
\end{aligned}
$$

$$
f(X', X'') = \frac{\Gamma}{2} \sum \frac{\frac{X'^{\delta\alpha \max(\phi^2)}}{X''_{\Gamma^{3/2}}} \sum \epsilon \sqrt{AB + AE_q - \frac{E_pE_q}{2} - \left(BE_p + \frac{E_pE_q}{2}\right)}}{\int_R \phi\rho f(\vec{X}|S_k) \frac{1}{(2\pi)^{d/2}\sigma^d} * \frac{1}{P_k} \sum_{i=1}^{P_k} \exp\left[-\frac{(\vec{X} - \overrightarrow{W_{ki}})^T (\vec{X} - \overrightarrow{W_{ki}})}{2\sigma^2}\right]}
$$

Ummm... okay...

Equations

Keep It Simple

Do you really need all those equations?

- This is very instance-dependent!
- Depends on what you're discussing
- Depends on your audience

Sometimes you may need them

- Explain the variables and what they mean
- Give a "plain-text" description of it

If you don't need them, don't use them!

Results

You have lots of cool results

No one can read this

No one can understand this

Graphs are your friend…

A	B	C	D	E
0.78799174	0.87677244	0.99348605	0.23781547	0.24437526
0.24910355	0.79708654	0.39825661	0.4894876	0.22079456
0.65729261	0.46901063	0.36471191	0.04697233	0.63468059
0.48205396	0.52657506	0.70503426	0.35280176	0.40935313
0.46328137	0.0774365	0.71517444	0.9394662	0.46843638
0.09762717	0.70884867	0.81407539	0.24571711	0.72497819
0.00773315	0.39906447	0.42344939	0.90776976	0.22209006
0.15857663	0.4181197	0.56488165	0.91405841	0.3578349
0.59242455	0.17894389	0.61926672	0.02978346	0.50789172
0.41285757	0.71470398	0.31906988	0.79658426	0.21587647
0.8855586	0.46534556	0.3701164	0.12452538	0.33415497
0.28231467	0.17509894	0.85801024	0.72984635	0.94731238
0.82370951	0.03235362	0.95622299	0.27726297	0.76619879
0.86245578	0.21094811	0.93272287	0.48265505	0.04960646
0.38953201	0.3665743	0.33754918	0.28178635	0.39637009
0.80522838	0.63509032	0.43333321	0.97677807	0.96198172
0.35928212	0.14878634	0.44201417	0.23251612	0.83375154
0.72099806	0.75212293	0.81061259	0.23756284	0.48518996
0.13329065	0.31602317	0.87489249	0.5304632	0.26191565
0.2588109	0.89039838	0.81380512	0.59139955	0.48488759
0.99314419	0.34635186	0.73292414	0.25933239	0.29230491
0.88041055	0.11473455	0.01934078	0.15717245	0.93780676
0.72332226	0.80195173	0.1792961	0.07832254	0.41154579
0.95925002	0.41696749	0.24905812	0.2111233	0.00256536
0.00580885	0.65322119	0.49666074	0.91641276	0.40573275
0.26004883	0.3010126	0.45604195	0.99935168	0.91271048
0.1508427	0.84418604	0.96241158	0.05548096	0.94093154
0.63750743	0.08979734	0.11100042	0.34646613	0.09994533
0.17176871	0.85518113	0.94522781	0.29368901	0.77444161
0.15186964	0.53105474	0.69991523	0.07876247	0.0023978
0.72306385	0.73755246	0.71402806	0.68090612	0.76015636
0.42140074	0.39036871	0.02247591	0.94725973	0.70692042

Visual aids

Legible and simple

Size at least 24-type

Resist seduction by color

Limit the number of equations

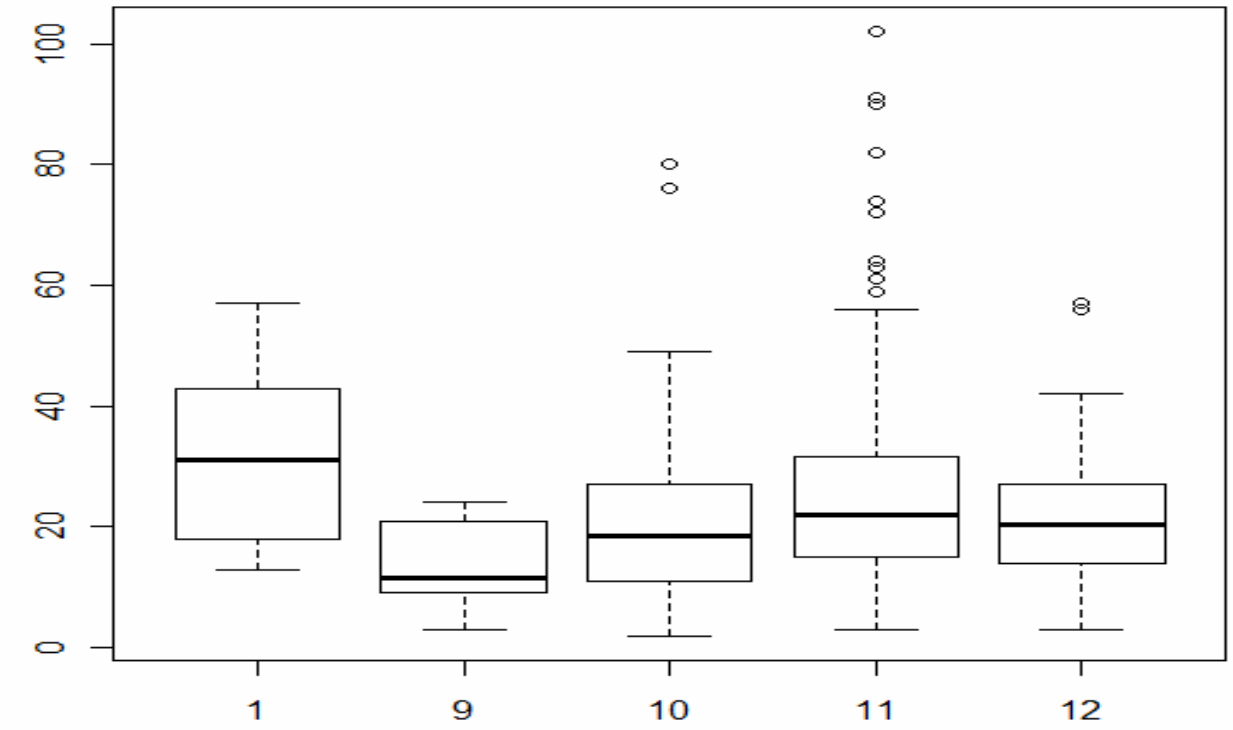
Limit the amount of animation

!! Don’t let your slides replace you !!

Visual aids

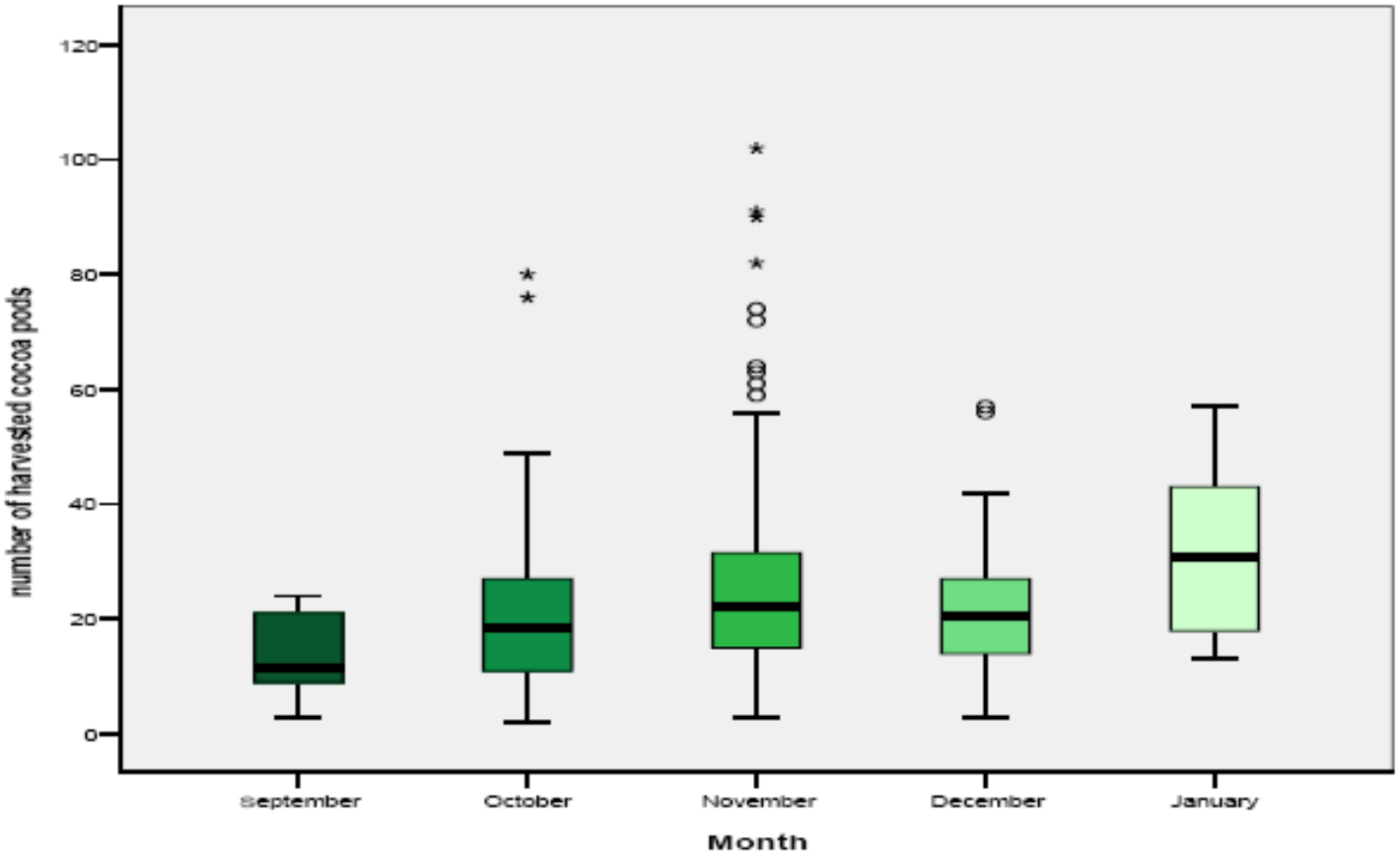
Always have a plan B. In case it might happen that the projector's bulb blows up in the course of the presentation, PowerPoint presentation or flip charts could be used

Bad graphs

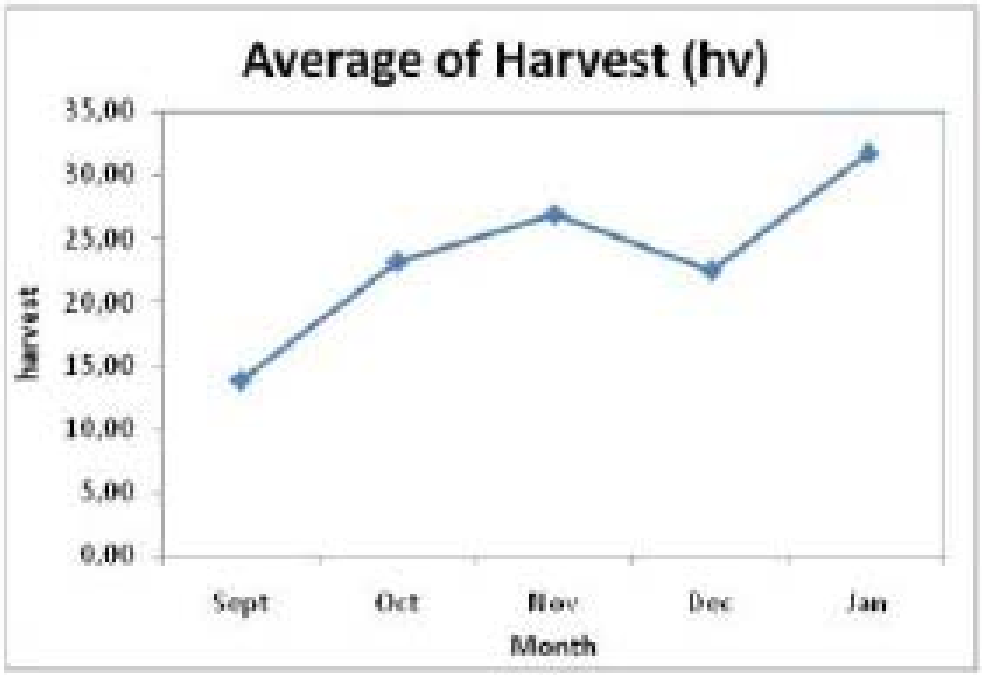


Group 1,3, 7, 11

Bad graphs

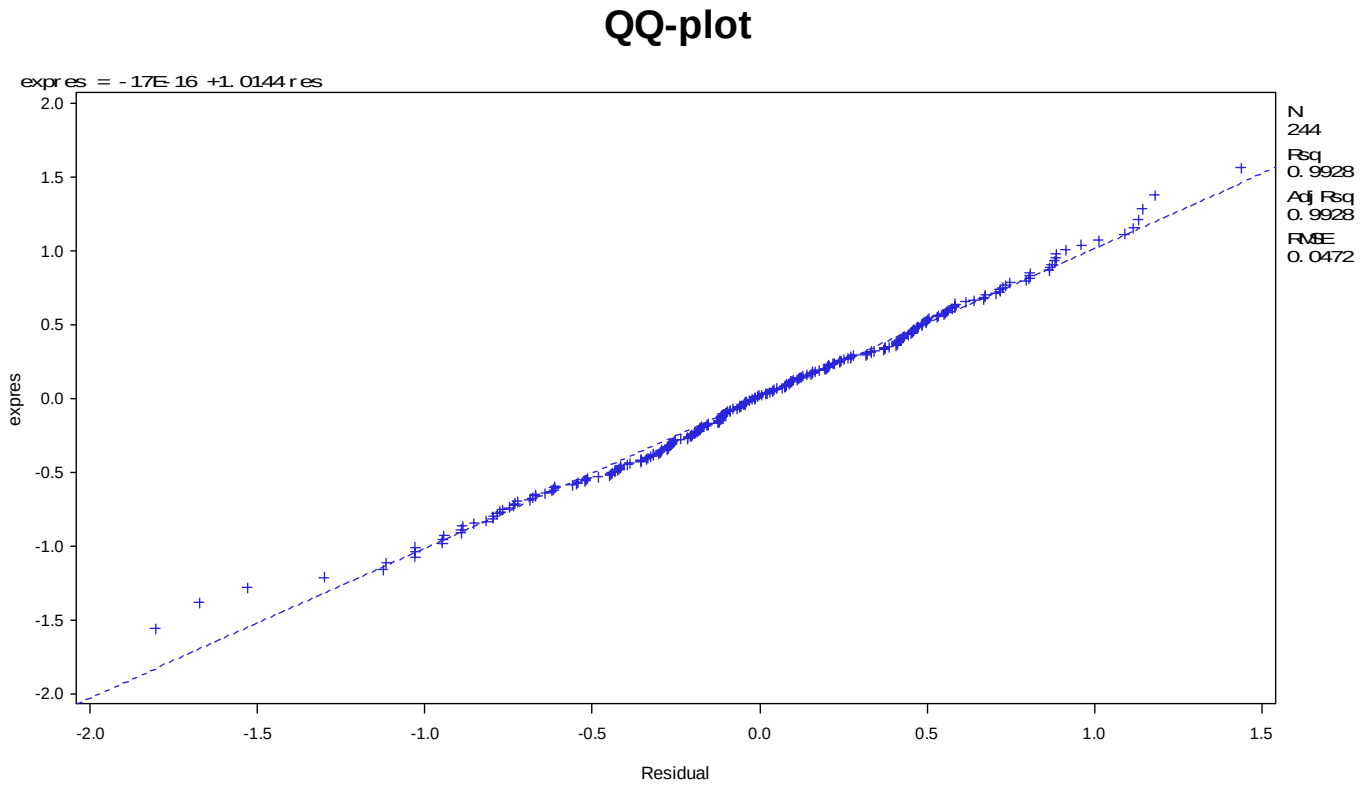

number of harvested cocoa pods
120
100
80
60
40
20
0
September
October
November
December
January
Month

Bad graphs



Graph A

Bad graphs


QQ-plot
expres = -17E-16 +1.0144 res
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
expres
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Residual
N
244
Rsq
0.9928
Adj Rsq
0.9928
RMSE
0.0472

Bad graphs

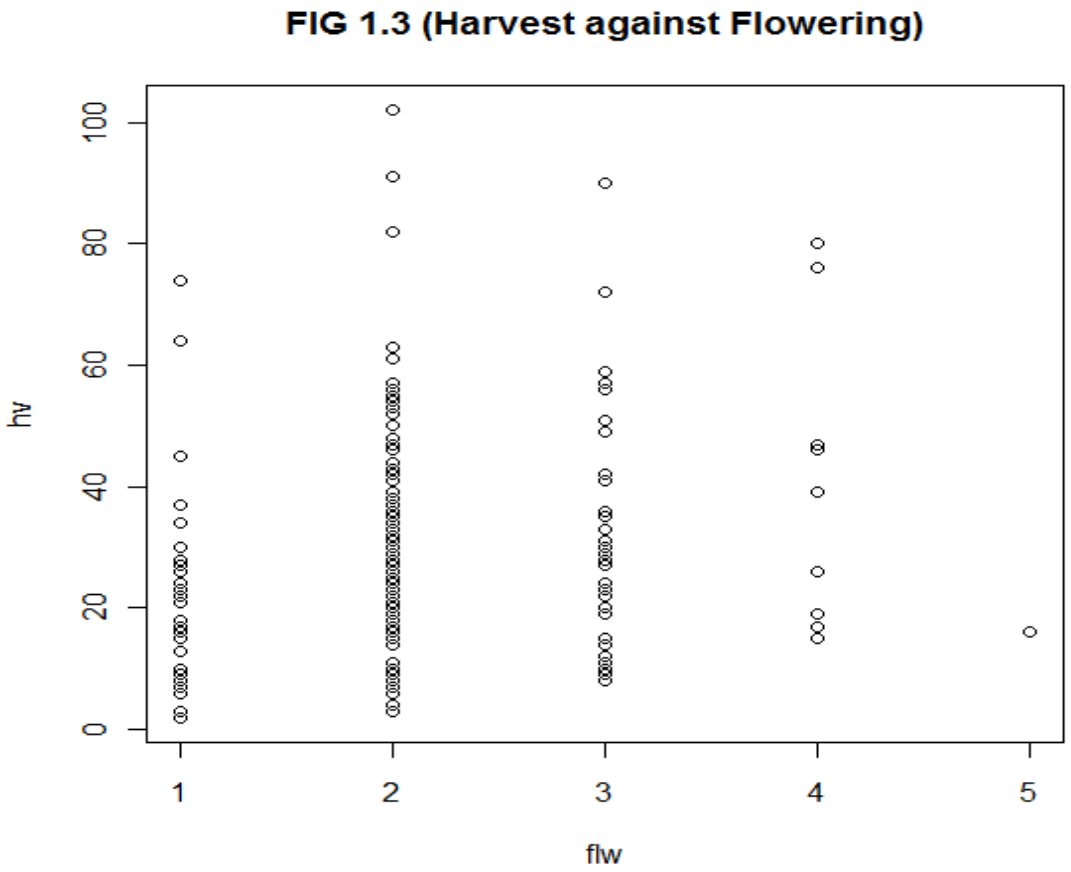
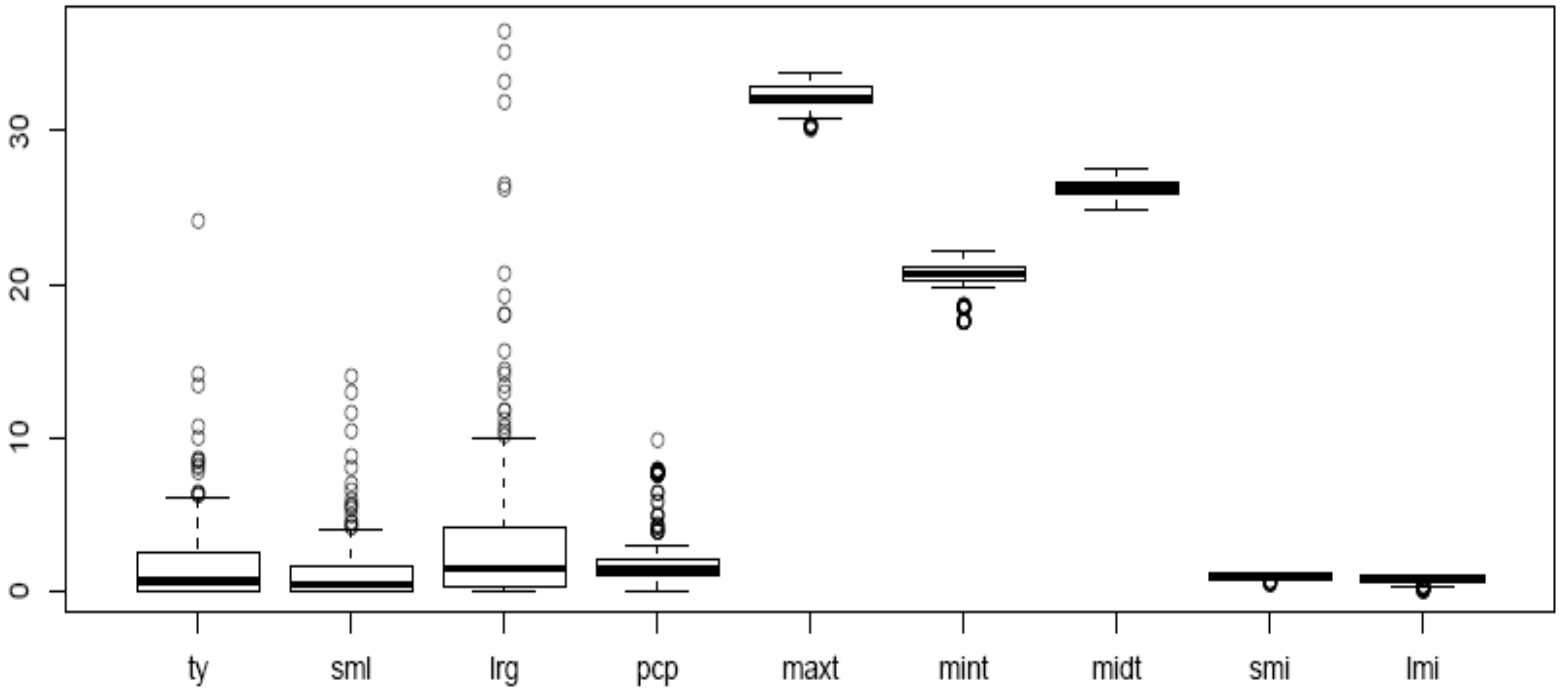


FIG 1.3 (Harvest against Flowering)
hv
0
20
40
60
80
100
1
2
3
4
5
flw

Bad graphs



Bad graphs

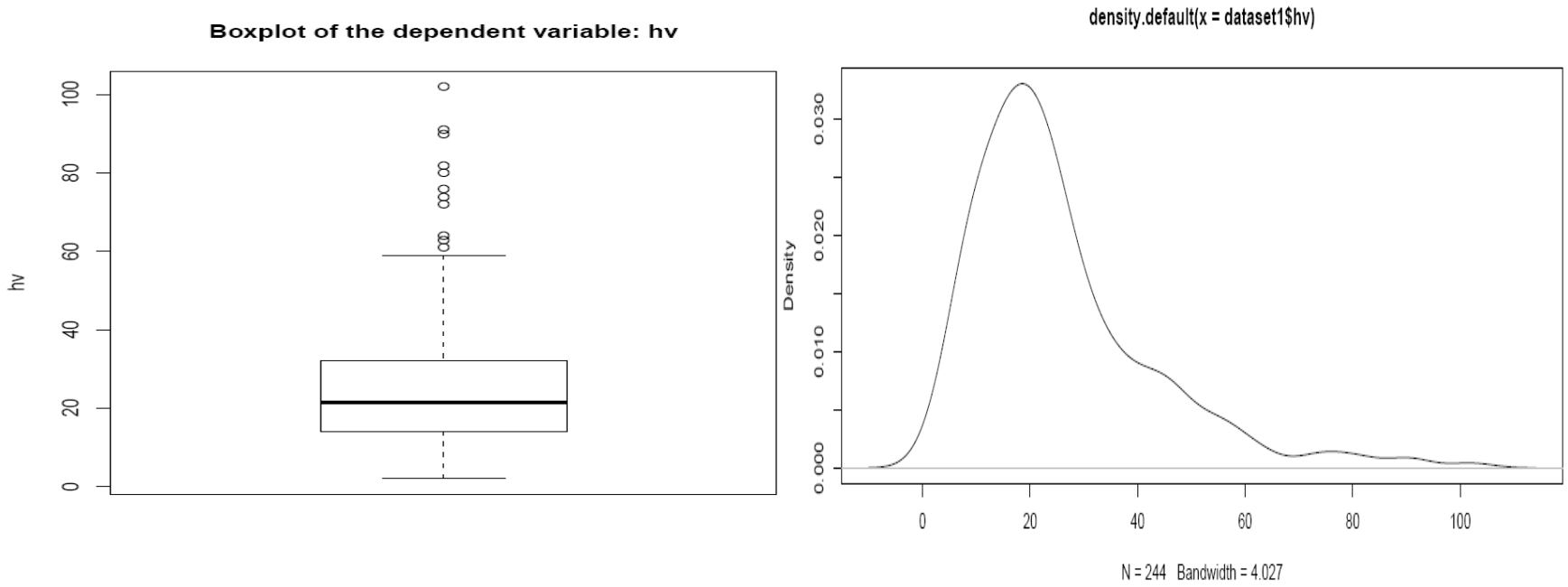


Figure 4.1 Box plot of dipendent variableand density function.

Bad graphs

FIG 2.1 (Harvest against Blk)

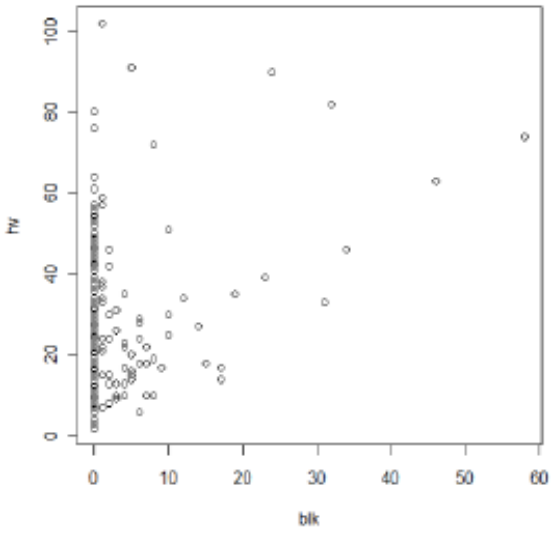


FIG 2.2 (Harvest against Wilt)

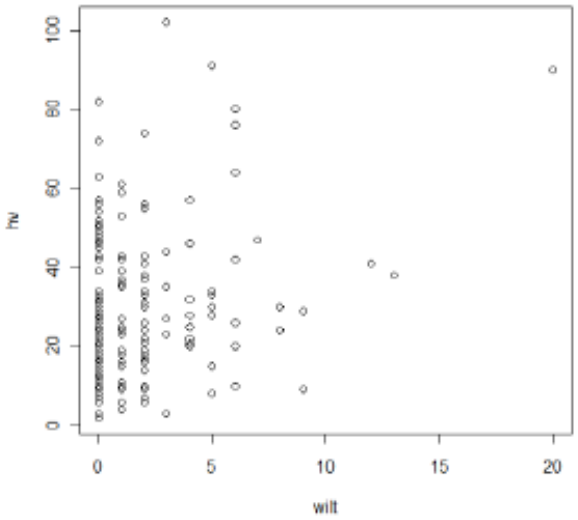
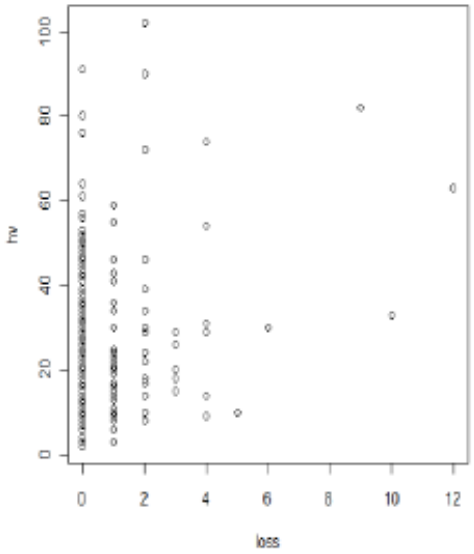
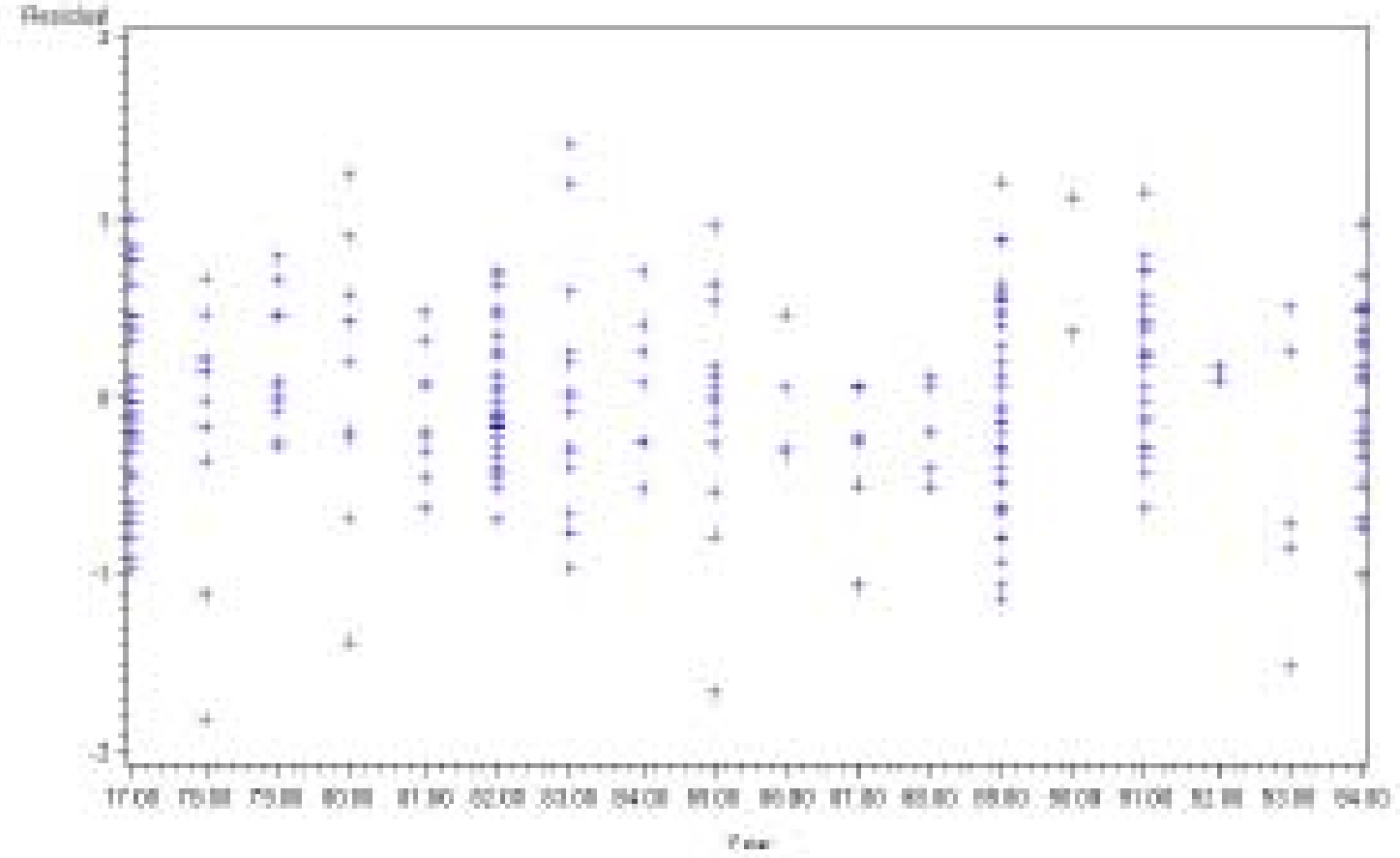


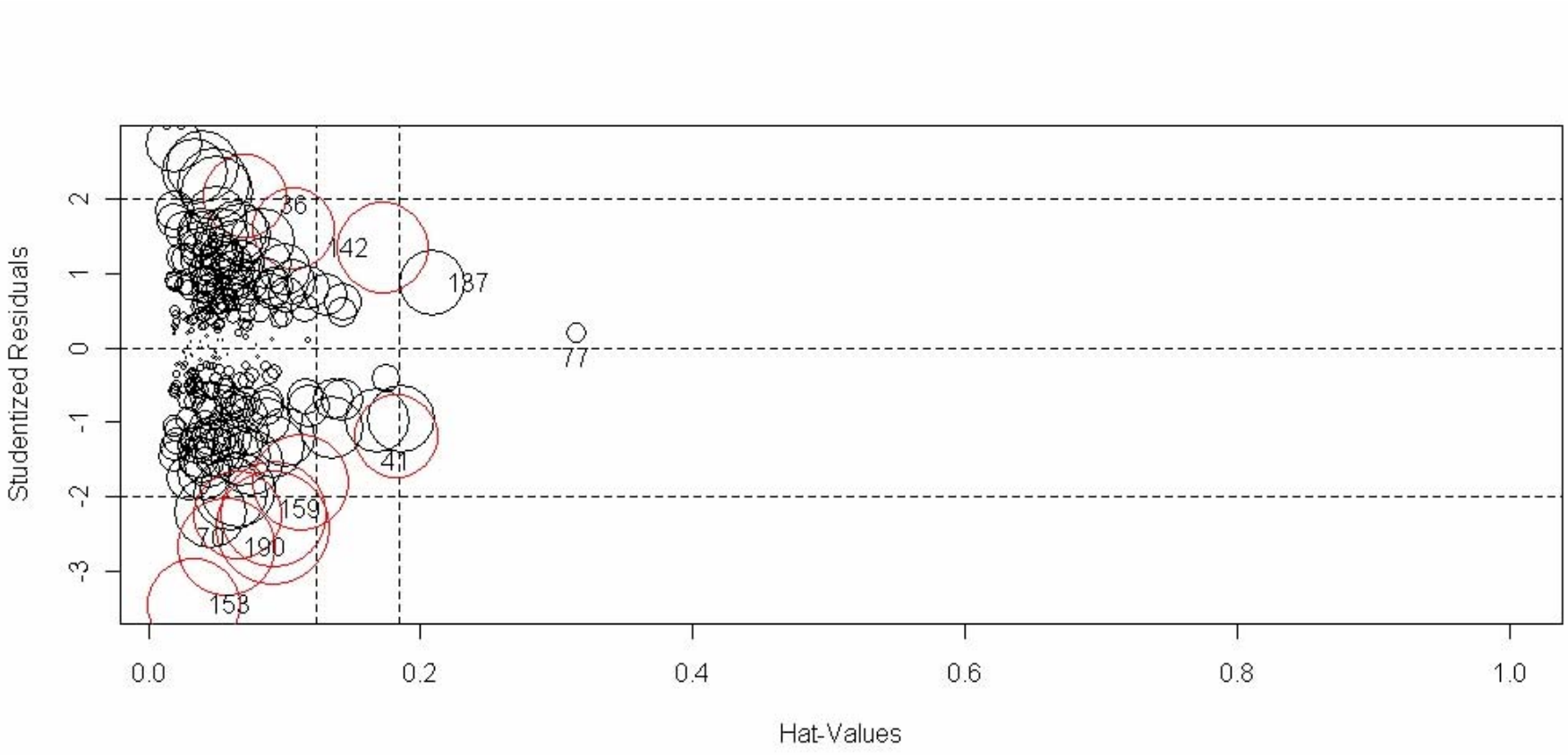
FIG 2.3 (Harvest against Loss)



Bad graphs



Bad graphs


Studentized Residuals
Hat-Values
0.0
0.2
0.4
0.6
0.8
1.0
2
1
0
-1
-2
-3
36
142
187
77
41
159
70
190
153

Giving criticism

- ✓ Be aware of emotional impact
- ✓ Be tactful
- ✓ Use assertive, non-aggressive communication
- ✓ Avoid *you*-based comments

Giving criticism

- ✓ Identify the positive
  - Start and finish with (2) positive comments
  - Negative is easy,
- ✓ Positive shows talent and creativity!

Giving criticism

- ✓ Depersonalize the message
  - Control your emotional state
- ✓ Provide possible solutions
- ✓ Focus on present and future
  - Don't dwell on the past
- ✓ Learn from your experiences!

Receiving criticism

- ✓ Criticisms are “easily” accepted if
  - well intended
  - effectively delivered
- ✓ Generally not a pleasant experience!

Receiving criticism

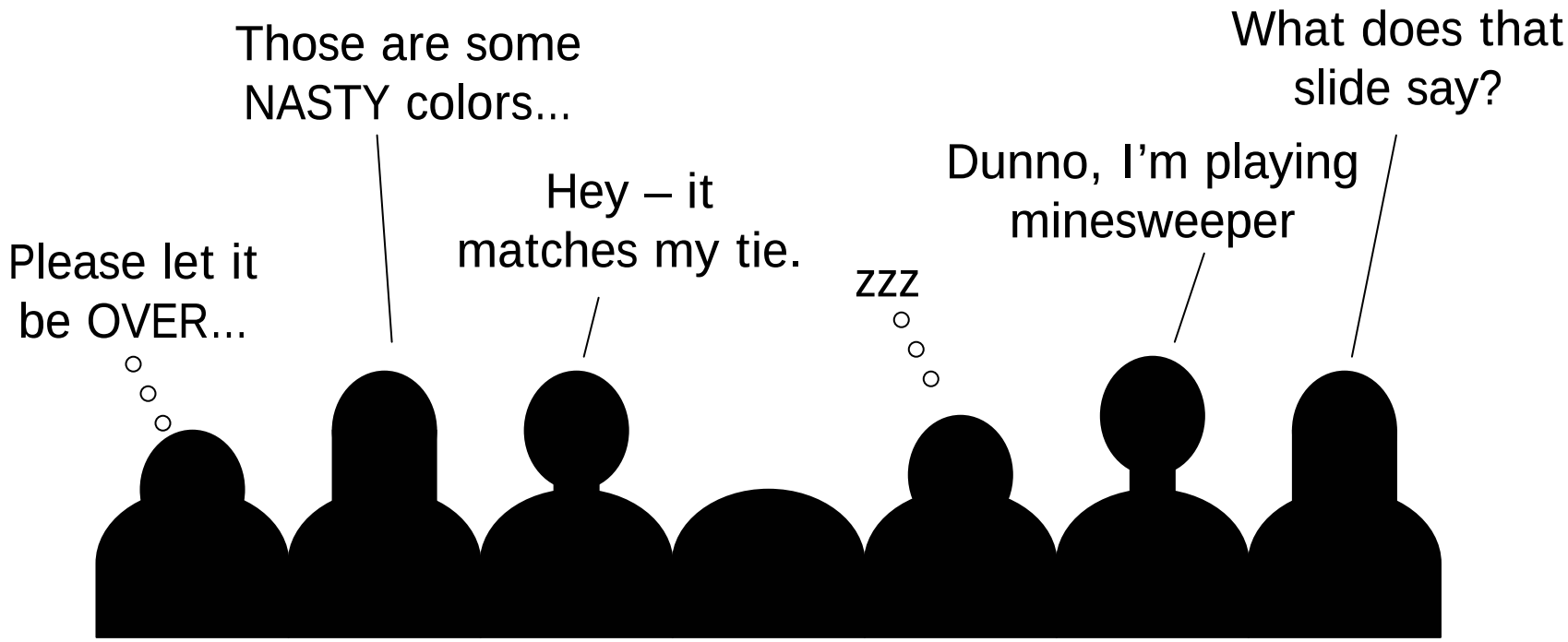
- ✓ Importance of honest feedback
  - Nobody is perfect
- ✓ Appreciate the value of having your work reviewed
- ✓ Don't take it personal!
- ✓ Clarify
  - Specific examples?
- ✓ Acknowledge
  - Identify
- ✓ Remedial steps?
  - Establish criteria
- ✓ Measure of "success"?

"Positive feedback always feels better and affirms your worth. Negative feedback properly received increases your value."

Chambers (2001)

Bad presentations

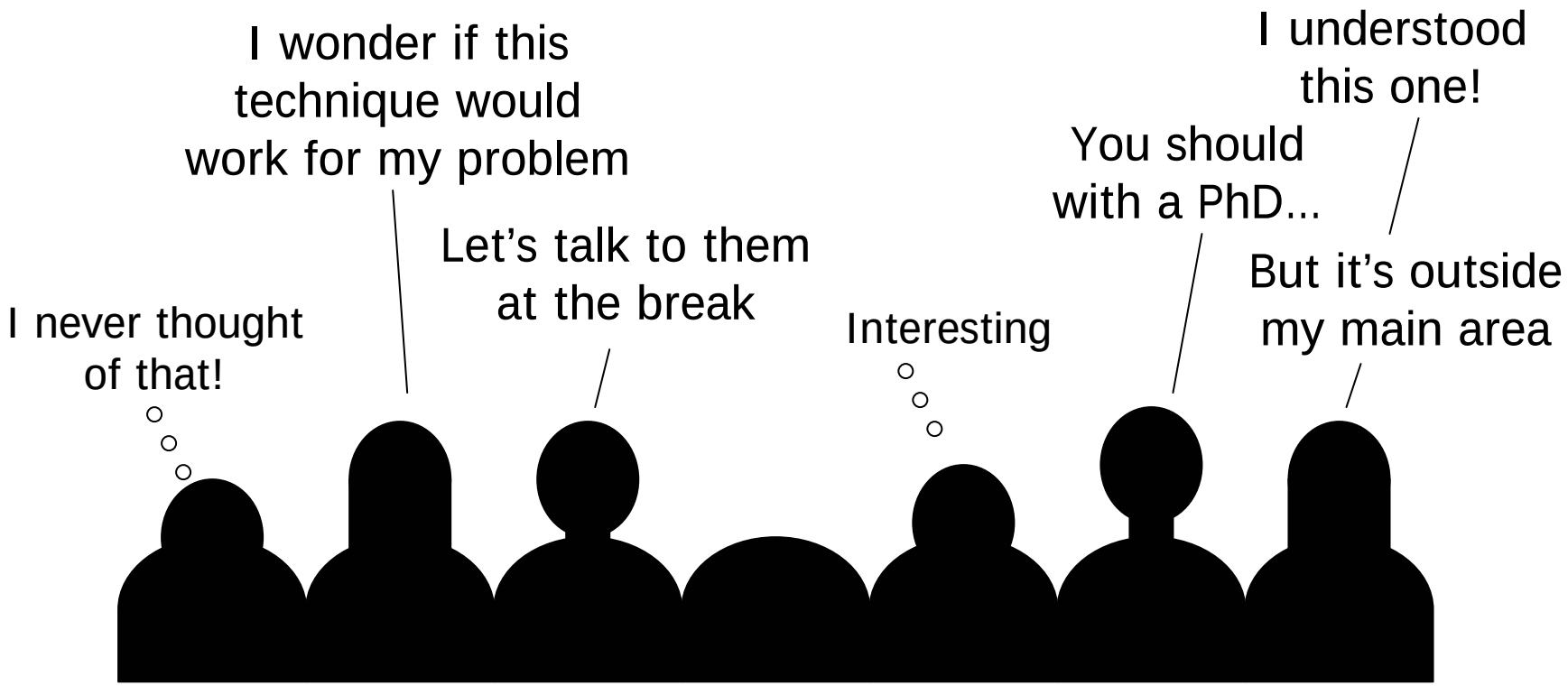
Audience won't see your work is great
But will make fun of you from back row



Good presentations

Interesting topic, explained at audience's level

Slides are understandable and easy to see



Some advice

"Oral presentation to many people is quite a nail-biting and nervous exercise. However, to some it is a joyous experience as they get to impact on the audience with information, and perfect the art of the skill"

Anonymous 2006-2007

Communication and Presentation

Written reports

Saskia Litière & Monique Maelstaf
Center for Statistics Hasselt University

Overview

Criteria for a written report

- ✓ Clear goal, precise and concise
- ✓ Constraints in language and format
- ✓ Structure
  - Title, short and strong
  - Introduction, what, why,how
  - Methods
  - Results
  - Discussion and conclusion

Use of visual aids

- ✓ References
- ✓ Appendix
- ✓ Header/footer

Visual aids

Use of visual aids (illustrations)

- Graphs
- Tables
- Figures
- Formulas

Use of figures and graphs

- when words are not sufficient
- for verisimilitude (proof of research findings)
- to emphasise information
- to provide background information
- to provide tools for critical analysis

Don't forget labels and titles!

Visual aids

Use of tables

- ✓ to summarize research findings
- ✓ to group together data sets
- ✓ to document experimental procedures and results
- ✓ to enable to make calculations from experimental data
- ✓ to enable reproducibility

Visual aids

Lack of Fit Test

This test will formally check whether the relation between the predictor and the response variables is linear.

The null and alternative hypothesis for this test are given by:

H_0:  ,

H_1:  .

To check whether the null hypothesis can be rejected the F test statistic is used:

$$F = \frac{\frac{SSLF}{c-p}}{\frac{SSPE}{N-c}} = \frac{MSLF}{MSPE}.$$

Visual aids

Since the observed studentized deleted residual is greater than 2 for observations 7, 35, 60, 80, 115, 119, 153, 119, 171 and 190, they are identified as outliers with respect to Y with observation 60 having the largest value of 3.08. Using the Bonferroni test procedure, the observed largest value is less than the tabulated critical value of $t\left(1-\alpha/2n;n-p-1\right)=t(0.9998;229)=3.59308$, hence, the outliers were found to be not influential with respect to the response.

Evaluating the hat diagonal H matrix, it was found out that 30 observations, as depicted in figure 5, were found to be outliers with respect to the predictors. These are observations number 1,4,19,25,34,41,45,61,74,77,79,80,89,95,112,113,121,122,123,135,137,141,142,155,159,172,176, 191,215 and 242 having h_{jj} values greater than the leverage value of 0.1148. Observation 41 was found to have the highest h_{jj} of 0.7783. Observation 80 was found to be an outlying observation with respect to both the predictors and the response

Visual aids

4.2 Model Selection

The first step in arriving at a final model, is to select the variables that are important to explain the response hv. As mentioned in the exploratory part of this report, there are some categorical variables. A consequence of this is that, in order to apply regression procedures, DUMMY-variables have to be created. The way this was done can be seen in the appendix. The method that was mentioned in the preceding subsection will become more clear by taking a look at the following code from SAS.

```
proc reg data=dummycocoa outest=selection noprint;
     model hv = jan sep okt nov  blk wilt loss can1 can2 can3 flu1 flu2 flu3 flw1 flw2 flw3 maxt mint pcp smi lmi ty sml lrg
     /selection=adjrsq  best=70000  aic adjrsq;
     run;
     quit;
```

```
data model;
set selection;
if jan= '.' & sep='.' & okt='.' & nov='.' then ;
else if jan^= '.' & sep^='.' & okt^='.' & nov^='.' then; else delete;
if can1= '.' & can2='.' & can3='.' then ;
else if can1^= '.' & can2^='.' & can3^='.' then ; else delete;
if flu1= '.' & flu2='.' & flu3='.' then ;
else if flu1^= '.' & flu2^='.' & flu3^='.' then ; else delete;
if flw1= '.' & flw2='.' & flw3='.' then ;
else if flw1^= '.' & flw2^='.' & flw3^='.' then ; else delete;
run;
```

Visual aids

We start by analyzing the effect of the different diseases on the harvest. The loss of cocoa pods could be due to black pods (Blk), Cherelle Wilt (Wilt) or due to some other diseases. From the graph (Appendix: FIG 2.1 – 2.3), we observe that as the level of each disease increases, the harvest decreases. But the losses due to black pod and Cherelle wilt have much effect on the total number of pods harvested than losses due to other infections. This could also be seen by regressing the harvest against all the diseases (Appendix: Output 1.).

Next we considered the effects of the Temperatures on the harvest. From the graphs (Appendix FIG 5.1- 5.3), we observe that the higher the temperature the better the harvest, the harvest is relatively constant for medium temperature, and for lower temperature the production is low. Consequently, we could expect that maximum and minimum temperatures should respectively have positive and negative effects on the production.

As can be seen from FIG 4.1-4.3(Appendix), as the amount of soil and leaf moisture indexes increase, the harvest is relatively good; and the harvest is good for less rainfall. As we know that there is less rainfall in the month of January and much in September (where the harvest is respectively highest and lowest) in Ivory Coast.

Visual aids

7. Appendix

Appendix A1. Descriptive Statistics

Statistics

		Hv	blk	wilt	loss	Ty	Sml	Lrg	Pcp	MaxT	MinT	MidT	Smi	Lmi
N	Valid	244	244	244	244	244	244	244	244	244	244	244	244	244
	Missing	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean		25,59	2,31	1,28	,61	1,7301	1,2151	3,6305	2,2107	32,1683	20,4073	26,2878	,86351	,6985
Median		21,50	,00	,00	,00	,6333	,3667	1,4500	1,3935	32,1300	20,6967	26,2450	,90994	,7597
Std. Deviation		17,097	6,790	2,450	1,485	2,79047	2,11795	5,94621	2,34231	,77857	1,05486	,63091	,110755	,2157

This means that the following model is used:

$$\ln(hv) = \beta_0 + \beta_1 cmaxt + \beta_2 cmint + \beta_3 cpcp + \beta_4 cblk + \beta_5 clrg + \beta_6 flw1 + \beta_7 flw2 + \beta_8 flw3 + \beta_9 cmaxt \cdot cpc$$

Visual aids

3.3. EXPLORATORY DATA ANALYSIS

Exploratory data analysis (EDA) was introduced by John Tukey as an approach to analyze data when there is only a low level of knowledge about its cause system as well as *contextual* information. EDA aims at letting the data itself influence the process of suggesting hypotheses instead of only using it to evaluate given *(a priori)* hypotheses. Explorative, opposed to Confirmatory data analysis is like detective work looking for patterns, analomies or in general new insights and is usually done via graphical representations of the underlying data-set.

Visual aids

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
hv	hv	244	25.59	17.10	2.00	102.00
blk	blk	244	2.31	6.79	0.00	58.00
wilt	wilt	244	1.28	2.45	0.00	20.00
loss	loss	244	0.61	1.49	0.00	12.00
ty	ty	244	1.80	2.80	0.00	24.00
sml	sml	244	1.27	2.16	0.00	14.00
lrg	lrg	244	3.67	5.96	0.00	37.00
pcp	pcp	244	2.21	2.34	0.00	9.88
maxt	maxt	244	32.17	0.78	30.20	33.85
mint	mint	244	20.41	1.05	17.61	22.24
midt	midt	244	26.29	0.63	24.91	27.55
smi	smi	244	0.86	0.11	0.51	0.99
lmi	lmi	244	0.70	0.22	0.08	0.96

Table 2

Visual aids

Table 3.3.2 Analysis of variance

Source	DF	SS	MS	F Value	Pr > F
Model	9	9.99633	1.11070	12.70	<.0001
Error	234	20.46396	0.08745		
Corrected Total	243	30.46029			

Parameter Estimates for the model

R-Square=0.3282 Adj R-Sq = 0.3023

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept		1	2.17914	0.06691	32.57	<.0001	0
BLK		1	0.00913	0.00294	3.10	0.0021	1.10793
LRG		1	0.02296	0.00366	6.28	<.0001	1.31858
FLW2		1	0.15922	0.04915	3.24	0.0014	1.68120
FLW3		1	0.15240	0.06389	2.39	0.0179	1.71311
FLW4		1	0.25016	0.10086	2.48	0.0138	1.22193
MONTH9		1	-0.48315	0.11529	-4.19	<.0001	1.45771
MONTH10		1	-0.44697	0.08090	-5.52	<.0001	2.29684
MONTH11		1	-0.20932	0.06598	-3.17	0.0017	3.00181
MONTH12		1	- 0.26371	0.07783	-3.39	0.0008	2.22234

Visual aids

Table 3: summary statistics for numeric explanatory variables

Variable	Mean	Median	N	Minimum	Maximum	Range
X_{blk}	2.3073770	0	244	0	58.0000000	58.0000000
X_{wilt}	1.2786885	0	244	0	20.0000000	20.0000000
X_{loss}	0.6065574	0	244	0	12.0000000	12.0000000
X_{ty}	1.7301230	0.6333333	244	0	24.2000000	24.2000000
X_{sml}	1.2150956	0.3666667	244	0	14.0000000	14.0000000
X_{lrg}	3.6305328	1.4500000	244	0	36.6000000	36.6000000
X_{pcp}	2.2107373	1.3935484	244	0	9.8774194	9.8774194
X_{maxt}	32.1683453	32.1299998	244	30.1999998	33.8500000	3.6500002
X_{mint}	20.4073258	20.6966667	244	17.6103448	22.2413793	4.6310345
X_{midt}	26.2878356	26.2450000	244	24.9089286	27.5483334	2.6394047
X_{smi}	0.8635104	0.9099372	244	0.5071888	0.9882381	0.4810493
X_{lmi}	0.6985927	0.7597859	244	0.0787578	0.9631328	0.8843750

Visual aids

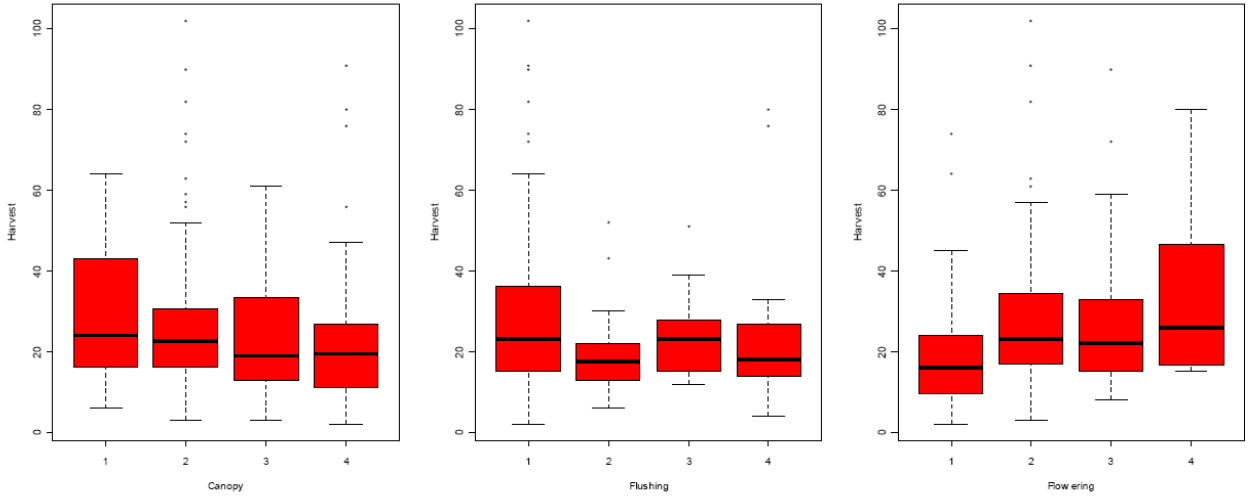


Figure 2: Box plots of the three categorical variables showing variation of different levels with mean monthly harvest for each variable.

Flowering:

The highest mean harvest (36.0) was observed at flowering stage 4, while flowering stage 1 recorded the lowest mean harvest. The harvest seemed to be higher for higher levels of flowering, indicating a positive relationship between the ordered flowering levels and the harvest.

Flushing:
The mean harvest was at its highest during the start of the flushing period (stage 1), and the lowest mean harvest is recorded in stage 2 of the flushing.
Canopy:

Visual aids

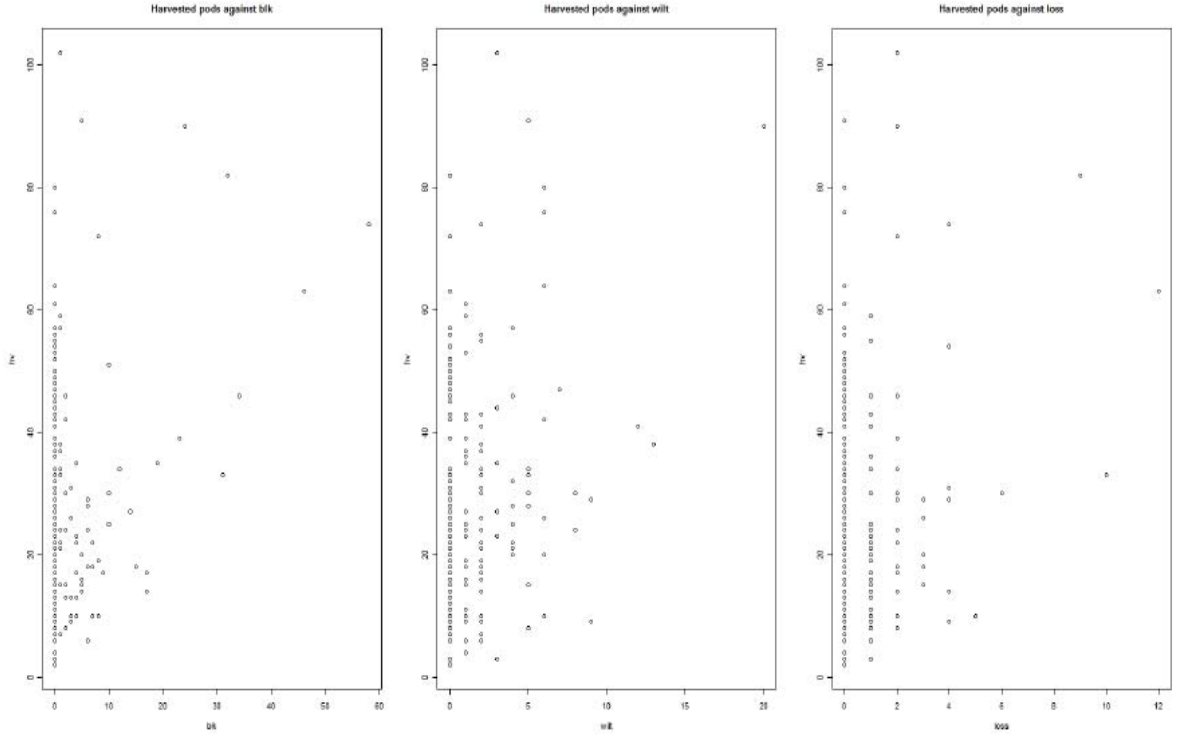


Figure 1.3: Harvested pods against diseases

References

References

Avoid plagiarism

Plagiarism

"is the deliberate attempt to deceive the reader through the appropriation and representation the work and words of others as one's own"

References

Plagiarism = stolen words =very serious offense

In Education and Examination Regulations
Sanctions:

- ✓failing grade on the particular assignment
- ✓ failing grade for the course
- ✓repeated plagiarism: student may be suspended or expelled

References

Essential proof of fair play in a scientists' world

The problem is not using the sources but failing to credit them

Direct quotations

- ✓ Between quotation marks
- ✓ Citations
- ✓ Footnotes

Indirect quotations

- ✓ Paraphrasing
- ✓ Synthesising (multiple sources)

References

Exercise

Paraphrasing

Like other exercise, laughter may have lasting benefits. Once the laughter stops, the muscles are more relaxed than before it started, which could relieve some kinds of headaches. Heartbeat and blood pressure dip to below normal, which is a sign of reduced stress. Laughter may be related to longevity, through this reduction of stress and hypertension.

Hamp-Lyons L. And Courter K.B. Research Matters, 1984, Oxford, Newbury House Publishers

- ✓ List significant, published references
- ✓ Check references before publication
- ✓ E-references

References

- ✓ Listed in alphabetical order
- ✓ Name, publishing date, title book, publisher:

Molenberghs, G. and Verbeke, G. (2008). *Models for Discrete Longitudinal Data*. New York: Springer

- ✓ Name, publishing date, title article, journal (volume), pages:

Aitkin, M. (2007). A general maximum likelihood analysis of variance components in generalized linear models. Biometrics 55, 117–128.

References

Citations: use "*et al.*" when 3 or more authors are involved

In text:

Molenberghs *et al*. (2007) used …

… (Aitkin, 2007; Aerts *et al*., 2005)

References

- ✓ E-References: URL's may fail
  - incorrectly typed
  - document moved or deleted

  Add: "Retrieved day, month, year"

  Add: Name of document, of institution.

University of Hasselt Library and Information Service, 2009, Bibliographic References Harvard Style. Diepenbeek: Universiteit Hasselt. Available at http:/www.uhasselt.be/bib/docs.html

(accessed 19/02/2010)

Final advice

Use header/footer

- ✓ Academic year
- ✓ Title project
- ✓ Number group

Appendix (superfluous??)

Summarizing

Exercise

Paper 1 “Probabilistic analysis of probe reliability in differential gene expression studies with short oligonucleotide arrays ”

10 minutes preparation

Group1 : Section “Modeling of Probe Reliability”

Group2: Section “Methods”

Result: a 60 sec. summary of your report including

- What
- Why
- How
- Conclusion

Working plan

Tasks division: who does what?

Time division:

Check advises from Special Topics slides lectures of 28th April SS2010

keywords.:

What?

How?

Exercise

Write a summary of your report

- Max 1 pages (no appendix!)
- Font choice
  - Word:Times New Roman, size 12
  - Latex, standard .12
- Double spaced
- Margins of 2cm

Conclusion

! Practice !
! Practice !
! Practice !