5.1 *k*-Anonymity and and *l*-Diversity

5.1.1 Table of 20 Random Selected Entries

Name	Surname	E-Mail	PLZ	City	System	Semester	Points
Ali	Muhammad	muhammad.ali@students.unibe.ch	3270	Aarberg	Android	5	89
Bristle	Mirko David Tobias	mirko.bristle@students.unibe.ch	3312	Fraubrunnen	Android	1	93
Brunner	Michael Sebastian	michael.brunner@students.unibe.ch	3510	Konolfingen	MacOS	12	82
Brunner	Julien Pierre	julien.brunner@students.unibe.ch	3800	Interlaken	Windows	1	82
Desarzens	Stéphane Patrick	stephane.desarzens@students.unibe.ch	3792	Saanen	Windows	1	82
Drazyk	Manuel	manuel.drazyk@unifr.ch	3150	Schwarzenburg	Android	2	54
Gerig	Pascal Dominik	pascal.gerig@students.unibe.ch	2520	La Neuveville	MacOS	8	6
Kutzschebauch	Lennart Platon	lennart.kutzschebauch@students.unibe.ch	3177	Laupen	Android	11	71
Lerena	Leandro Patricio Kai	leandro.lerena@students.unibe.ch	3294	Büren	MacOS	1	27
Lüthi	Thomas	thomas.luethi1@students.unibe.ch	3380	Wangen	Android	7	23
MÉNÉTREY	Jämes	james.menetrey@unine.ch	3456	Trachselwald	Windows	5	45
Moudi	Jonas	jonas.moudi@students.unibe.ch	3600	Thun	Linux	6	18
Schmucker	Adrian Walter	adrian.schmucker@students.unibe.ch	3860	Oberhasli	Windows	12	31
Senn	Michael Andreas	michael.senn@students.unibe.ch	3792	Saanen	iOS	2	60
Vonlanthen	Jérôme David Julien	jerome.vonlanthen@unifr.ch	3534	Signau	Linux	4	13
Weyermann	Julian Dominic	julian.weyermann@students.unibe.ch	3456	Trachselwald	MacOS	10	11
Wipfli	Lorenzo	lorenzo.wipfli@students.unibe.ch	4912	Aarwangen	MacOS	10	33
Zauder	Marcel Matti	marcel.zauder@students.unibe.ch	3770	Obersimmental	MacOS	11	48
Zeiter	Sandrine	sandrine.zeiter@unifr.ch	3763	Niedersimmental	Android	9	31
Zuber	Joël Luca	joel.zuber@students.unibe.ch	3714	Frutigen	Windows	3	2

5.1.a Define identifiers, two quasi-identifying attributes, and one sensitive attribute.

Identifiers: Name, Surname and E-Mail adress.
Quasi-Identifying Attributes: PLZ and/or City (are dependent on each other), Points
Sensitive Attribute: System

5.1.b Sanitize this data by computing a 3-anonymized version of this data by hand.

I will define PLZ and Points-Ranges, in order to create clusters that share characteristics with each other, ensuring 3-Anonymity (see Table 5.1).

5.1.c Does it become easier or more difficult to derive a 5-anonymized data set?

It is easy to determine a 5-anonymized data set, because one can take the subsets from 5.1.b and merge them together such that the set containing the least entries does contain at least 5 items.

5.1.d Sanitize the entries to satisfy 3-diversity.

Using the previously created 3-Anonymity data set, by adjusting the ranges clusters can be created which contain at least 3 different System types, in order to ensure 3-Diversity (see Table 5.2).

PLZ-Range	Points-Range	System
2500-3200	0-100	MacOS
		Android
		Android
3200-5000	0-15	Windows
		MacOS
		Linux
	16-30	Linux
		Android
		MacOS
	31-50	Android
		Windows
		MacOS
		Windows
		MacOS
	50-100	iOS
		MacOS
		Windows
		Windows
		Android
		Android

Table 5.1: 3-Anonymity Dataset

PLZ-Range	Points-Range	System
2500-3750	0-15	Windows
		MacOS
		MacOS
		Android
	16-30	Linux
		Android
		MacOS
3150-3800	46-80	MacOS
		Android
		iOS
		Android
	81-100	MacOS
		Windows
		Windows
		Android
		Android
3450-5000	31-45	Android
		Windows
		MacOS
		Windows

Table 5.2: 3-Diversity Dataset