

<b>Chapter 1 – Projekt EverGreen .....</b>	<b>7</b>
<b>Chapter 2 – Uvod u znanost o podacima .....</b>	<b>9</b>
2.1 Koncept velikih skupova podataka i kako se razlikuje od tradicionalnih izvora podataka .....	9
2.2 Izazovi rada s velikim podacima .....	11
2.3 Struktura podataka i njezina važnost u znanosti o podacima .....	12
2.4 Strukturirani podaci u odnosu na nestrukturirane podatke: .....	13
2.5 Otkrivanje raznolikosti u zastupljenosti podataka.....	13
2.6 Tehnike predobrade podataka .....	14
2.6.1 Čišćenje podataka.....	14
2.6.2 Normalizacija podataka .....	14
2.7 Opcije spremišta podataka: relacijske baze podataka, NoSQL baze podataka i jezera podataka..	15
2.7.1 Relacijske baze podataka.....	15
2.7.2 NoSQL Baze podataka .....	15
2.7.3 Jezera podataka .....	15
2.8 Tehnike obrade i analize podataka: strojno učenje, obrada prirodnog jezika i vizualizaciju podataka .....	16
2.8.1 Strojno učenje.....	16
2.8.2 Obrada prirodnog jezika (NLP).....	16
2.8.3 Vizualizacija podataka .....	17
<b>Chapter 3 – Ciljevi održivog razvoja.....</b>	<b>19</b>
<b>Chapter 4 – Statistika i analiza podataka .....</b>	<b>23</b>
4.1 Mjere središnje tendencije.....	23
4.1.1 Srednja vrijednost.....	23
4.1.2 Medijan .....	23
4.1.3 Mod .....	24
4.2 Mjere disperzije.....	24
4.2.1 Raspon.....	24
4.2.2 Varijanca .....	24
4.2.3 Standardna devijacija .....	25
4.2.4 Standardna pogreška.....	25
4.3 Histogrami i pravokutni dijagrami (boxplotovi) .....	25
4.4 Testiranje hipoteze .....	27
4.4.1 Z - test .....	28
4.4.2 T-test .....	29
4.4.3 Hi-kvadrat test .....	29
4.5 Regresijska analiza.....	30
4.5.1 Linearna regresija .....	31
4.5.2 Višestruka linearna regresija .....	32
4.5.3 Korištenje višestruke linearne regresije u programu Excel .....	32
4.6 Klasifikacija i njezine primjene.....	34
4.6.1 Logistička regresija .....	34
<b>Chapter 5 – Statistika i analiza podataka .....</b>	<b>37</b>
5.1 Proračunske tablice (XLSX) .....	37
5.2 Vrijednosti odvojene zarezima (CSV) .....	38
5.3 Jezik za označavanje - eXtensible Markup Language (XML) .....	38
5.4 JavaScript Object Notation (JSON) .....	39
5.5 YAML.....	40
5.6 RESTful servisi .....	40
5.7 Digitalna arhiva videozapisa (DVA).....	41
5.8 Tok podataka.....	42
5.9 Tablice u bazama podataka .....	42
5.10 Vanjske (eksterne) tablice.....	43
5.11 Drugi formati podataka .....	44
<b>Chapter 6 – Alati za analizu podataka .....</b>	<b>47</b>
6.1 Microsoft Excel.....	47
6.2 Python .....	48

6.3	R.....	49
6.4	Tableau.....	50
6.5	Apache Spark .....	50
6.6	SAS .....	51
6.7	Power BI .....	51
6.8	Orange.....	52
6.9	JupyterLab.....	53
6.10	KNIME Analytics Platform .....	53
6.11	Oracle Analytics Cloud.....	54
<b>Chapter 7 – Analiza podataka u SQL-u .....</b>		<b>57</b>
7.1	Podrška za transakcije .....	57
7.2	Normalizacija podataka.....	58
7.3	Indeks .....	60
7.3.1	ROWID .....	60
7.3.2	Upravljanje indeksima.....	61
7.3.3	Vrste indeksa .....	61
7.3.4	Access metoda pristupa .....	64
7.4	Dohvaćanje podataka - odabir naredbe .....	64
7.5	Plan izvođenja .....	65
7.6	Funkcije agregiranja.....	67
7.6.1	Razmatranje jedinstvenih vrijednosti .....	70
7.6.2	Proširenja Rollup i Cube .....	70
7.6.3	Uvjeti koji se temelje na agregatnim funkcijama .....	72
7.6.4	Dohvaćanje vrijednosti na temelju agregatnih funkcija .....	73
7.6.5	Dohvati prvi upit .....	74
7.6.6	Sortiranje na temelju podataka koji nisu jedinstveni .....	76
7.7	Analitičke funkcije.....	78
7.7.1	Pravilo vodećih N.....	79
7.7.2	Filtriranje na temelju analitike.....	80
7.7.3	NULL vrijednosti i analitika .....	82
7.7.4	Analitika upravljanja vremenom .....	82
7.7.5	Nth_value .....	84
7.7.6	Skup podataka .....	86
7.7.7	Najam automobila .....	87
7.7.8	Praćenje temperature pomoću analitičkih funkcija Lead i Lag .....	88
7.7.9	Uvjeti proširene analitičke funkcije.....	92
7.7.10	Vanjske tablice .....	95
<b>Chapter 8 – Skladišta i tipovi podataka .....</b>		<b>99</b>
<b>Chapter 9 – Oracle Data Integrator .....</b>		<b>107</b>
<b>Chapter 10 – Uvod u Python .....</b>		<b>123</b>
10.1	Naredbe i funkcije.....	125
10.2	Python biblioteke za analizu podataka.....	129
10.2.1	NumPy biblioteka.....	129
10.2.2	Pandas .....	131
10.3	Vizualizacija podataka.....	144
10.3.1	Linijski grafikon .....	144
10.3.2	Stupčasti grafikon.....	146
10.3.3	Raspršeni grafikon.....	147
10.3.4	Tortni grafikon .....	147
10.3.5	Box grafikoni.....	150
10.3.6	Violin grafikon .....	151
10.3.7	Toplinske karte .....	152
10.3.8	Mjehurićasti grafikon .....	153
10.4	Strojno učenje u Pythonu.....	156
10.4.1	Linearna regresija.....	156
10.4.2	Neuronska mreža.....	158

10.4.3	Algoritam K srednjih vrijednosti .....	161
10.4.4	Algoritam K-najbliži susjeda (KNN) .....	162
<b>Chapter 11 – Analiza podataka u Sparku .....</b>		<b>167</b>
11.1	Hadoop.....	167
11.2	Grozdi za velike količina podataka.....	167
11.3	Spark i programski jezici .....	168
11.4	Osnovni principi rada s velikom količinom podataka (u Sparku).....	172
11.4.1	Osnovne Python biblioteke za analizu podataka .....	172
11.4.2	Pandas .....	173
11.4.3	DataFrames.....	173
11.4.4	DataFrame u Sparku .....	176
11.5	SQL u Sparku .....	178
11.6	NumPy .....	179
11.7	Matplotlib .....	180
11.8	SciPy.....	180
11.9	TensorFlow (i PyTorch, Keras).....	180
11.10	Vanjski izvori podataka .....	180
11.10.1	Csv datoteke.....	180
11.10.2	Oracle baza podataka .....	181
11.10.3	MS SQL poslužitelj.....	181
11.11	Paralelna obrada.....	181
11.11.1	RDD .....	181
11.11.2	Mapiraj i reduciraj (map and reduce).....	182
11.11.3	Transformacije - uglavnom funkcije mapiranja, ali i druge .....	182
11.11.4	Događaji.....	183
<b>Chapter 12 – Oracle Cloud.....</b>		<b>185</b>
12.1	SQL Developer povezivanje s bazom podataka .....	206
12.2	SQL*Plus – SQL Client.....	209
12.2.1	Puno definiranje povezivanja .....	210
12.2.2	Identifikatori povezivanja.....	211
12.2.3	Komentari.....	214
12.2.4	Povezivanje i prekid sesije .....	215
<b>Chapter 13 – Korištenje OCI Analytics Cloud platforme .....</b>		<b>217</b>
<b>Chapter 14 – Vizualizacija podataka.....</b>		<b>235</b>
14.1	Vizualizacija podataka.....	236
14.2	Oracle Analytics .....	237
<b>Chapter 15 – Studijski slučaj .....</b>		<b>245</b>
⇒ Studijski slučaj Općine Kranj ⇐ .....		245
15.1	Uvod .....	245
15.1.1	Kranj kao pametan grad .....	245
⇒ Studijski slučaja s bazom podataka World Ocean ⇐ .....		267
15.2	Uvod .....	267
15.2.1	Baza podataka World Ocean .....	267
15.2.2	Formati znanstvenih podataka.....	268
15.3	Priprema baze podataka za OAC platformu .....	268
15.3.1	Dohvaćanje podataka iz servisa Oracle Open Data .....	268
15.3.2	Izdvajanje i transformacija podataka u nc datotekama .....	269
15.3.3	Pohrana površinskih temperatura u OCI bazi podataka.....	269
15.3.4	Priprema početnog skupa podataka WOD podataka u OAC-u.....	272
⇒ Studijski slučaj - korištenje Kaggle otvorenih podataka o okolišu ⇐ .....		287
15.4	Istraživanje i korištenje otvorenih podataka o okolišu.....	287
15.5	Jednostavna integracija podataka.....	290
15.6	Osnove vizualizacije podataka.....	292
15.7	Osnovna analitika .....	294
15.8	Ponovimo što smo naučili u servisu Oracle Analytics .....	299

⇒ <b>Studijski slučaj: Upravljanje zaposlenicima</b> ⇐ .....	307
15.9 Upotreba OCI analitike .....	307
⇒ <b>Studijski slučaj: Praćenje temperature i predviđanje</b> ⇐ .....	321
15.10 Dohvat i učitavanje podataka .....	321
15.11 Izrada modela predviđanja temperature .....	329
15.12 Implementacija modela u Oracle APEX okolini .....	333
<b>Chapter 16 – Iskorištavanje podataka</b> .....	<b>353</b>
16.1 Prikupljanje podataka .....	353
16.2 Priprema podataka .....	354
16.3 Analiza podataka .....	354
16.4 Stvaranje uvida .....	355
16.5 Odlučivanje .....	355
16.6 Stalno poboljšanje .....	355
16.7 Etička razmatranja .....	356
16.8 Stvaranje vrijednosti .....	356
16.9 Budućnost eksploatacije podataka .....	356
16.9.1 Iskorištavanje podataka za održivost .....	357
16.9.2 Pejzaž zelene industrije .....	357
16.10 Budućnost iskorištavanja podataka u zelenoj industriji .....	357
16.11 Zaključak .....	358
<b>Chapter 17 – Strategija poslovnih podataka</b> .....	<b>359</b>
17.1 Usklađenost s poslovnim ciljevima .....	359
17.2 Upravljanje podacima .....	359
17.3 Prikupljanje i integracija podataka .....	359
17.4 Pohrana podataka i infrastruktura .....	360
17.5 Analitika podataka i uvidi .....	361
17.6 Upravljanje životnim ciklusom podataka .....	361
17.7 Sigurnost podataka i privatnost .....	362
17.8 Suradnja i dijeljenje podataka .....	362
17.9 Monetizacija podataka .....	363
17.10 Tehnologija i alati .....	364
17.11 Metrike i mjerenje učinka .....	365
17.12 Stalno poboljšanje .....	365
17.13 Prednosti dobre podatkovne strategije .....	366
17.13.1 Ključne komponente učinkovite podatkovne strategije .....	367
17.13.2 Različiti pristupi za izradu podatkovne strategije .....	367
17.13.3 Tko gradi strategiju poslovnih podataka? .....	368
17.14 Zaključci .....	368
<b>Chapter 18 – Izvrsnost podataka</b> .....	<b>369</b>
18.1 Kvaliteta .....	369
18.2 Relevantnost .....	369
18.3 Pristupačnost .....	369
18.4 Dosljednost .....	369
18.5 Pravovremenost .....	369
18.6 Sigurnost i privatnost .....	369
18.7 Upravljanje podacima .....	369
18.8 Integracija podataka .....	370
18.9 Analiza podataka i uvidi .....	370
18.10 Stalno poboljšanje .....	370
18.11 Prednosti izvrsnosti podataka .....	370
18.12 Načela izvrsnosti podataka .....	371
18.13 Kako izmjeriti i poboljšati izvrsnost podataka? .....	371
18.14 Izvrsnost podataka u zelenim podacima .....	372
18.15 Zaključak .....	374
<b>Chapter 19 – Zaključak</b> .....	<b>375</b>