

BIG DATA SYSTEMS

PROJECT 1: BIG MOBILITY & IOT DATA ANALYTICS

NADA JOVANOVIĆ 1618

ZADATAK

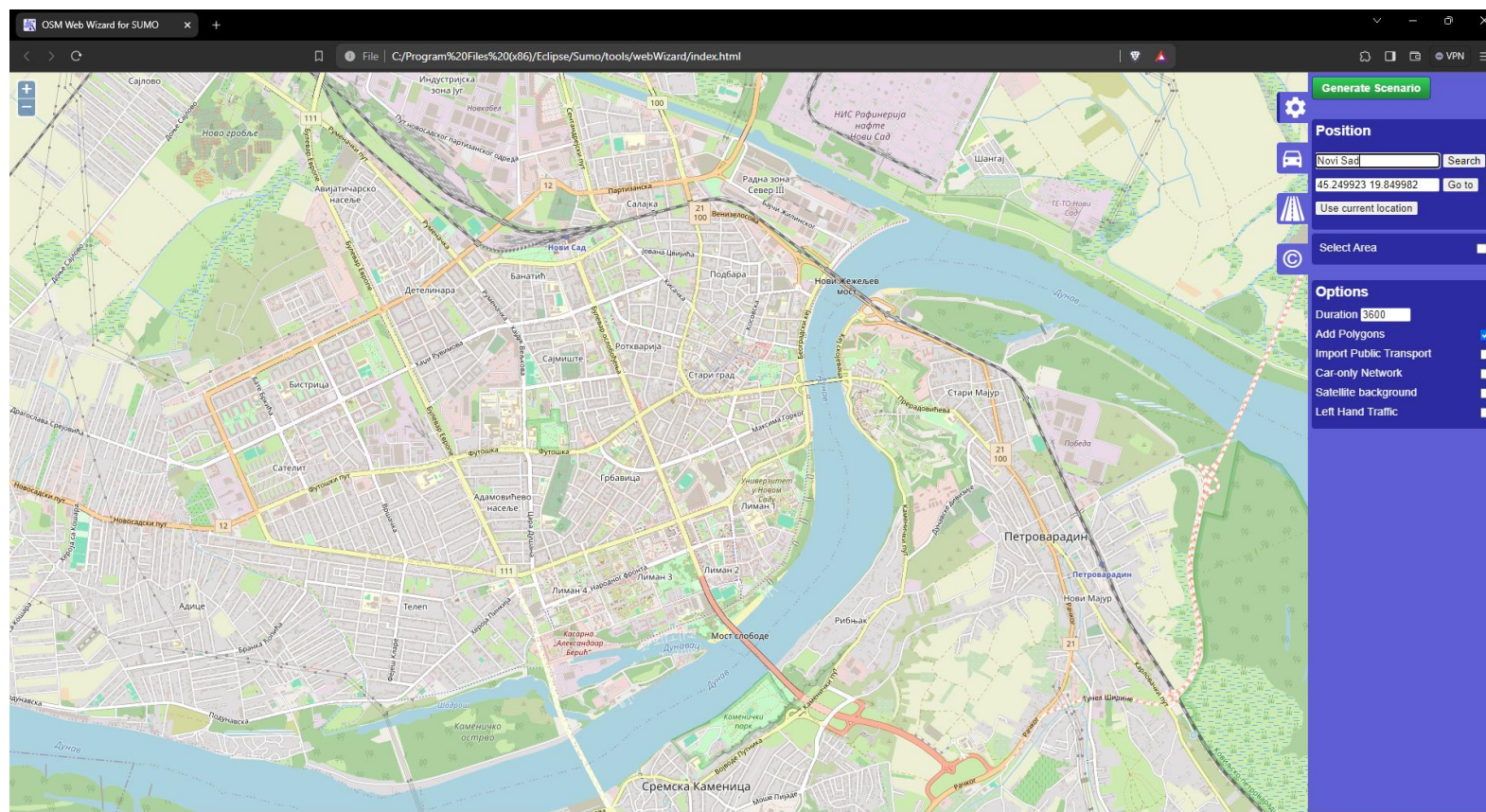
- Korišćenjem **SUMO** simulatora <https://sumo.dlr.de> kreirati jedan ili više velikih skupova podataka (~1GB) koji se odnose na kretanje vozila u IoT sistemima pametnih gradova i pametnog saobraćaja/transporta, u određenom gradu u Evropi ili svetu. Dobijene podatke u XML formatu prebaciti u .csv (xml2csv.py) i postaviti podatke na **HDFS**
- Razviti aplikaciju korišćenjem **Apache Spark** framework-a i Python (Scala, Java) API za batch obradu i analizu ovih podataka koja uključuje sledeće funkcionalnosti:
 1. Odrediti broj vozila i njihove karakteristike koja su prolazila u blizini značajnih mesta u gradu u određenom vremenskom periodu. Mesto (geo koordinate mesta), veličina „blizina“, granice vremenskog perioda i eventualni filter (npr. tip vozila) se zadaju kao argumenti aplikacije.
 2. Naći minimalne, maksimalne, srednje vrednosti, standardne devijaciju i ostale statističke parametre zagađenja (CO₂, CO, HC, NO_x, PM_x, noise) ili potrošnje goriva (fuel, electricity) po ulicama (edge/lanes) u određenim vremenskim periodima.

Korišćene tehnologije



GENERISANJE PODATAKA

GENERISANJE MREŽE – OSM WEB WIZARD









































GENERISANJE OUTPUT-A

- Za generisanje output-a korišćene su sledeće komande:
 - `sumo -c osm.sumocfg --fcd-output fdc.xml --fcd-output.geo`
 - `sumo -c osm.sumocfg --emission-output emissions.xml`
- A zatim:
 - `python tools/xml/xml2csv.py "D:\novisad\emissions.xml"`
 - `python tools/xml/xml2csv.py "D:\novisad\ fdc.xml"`

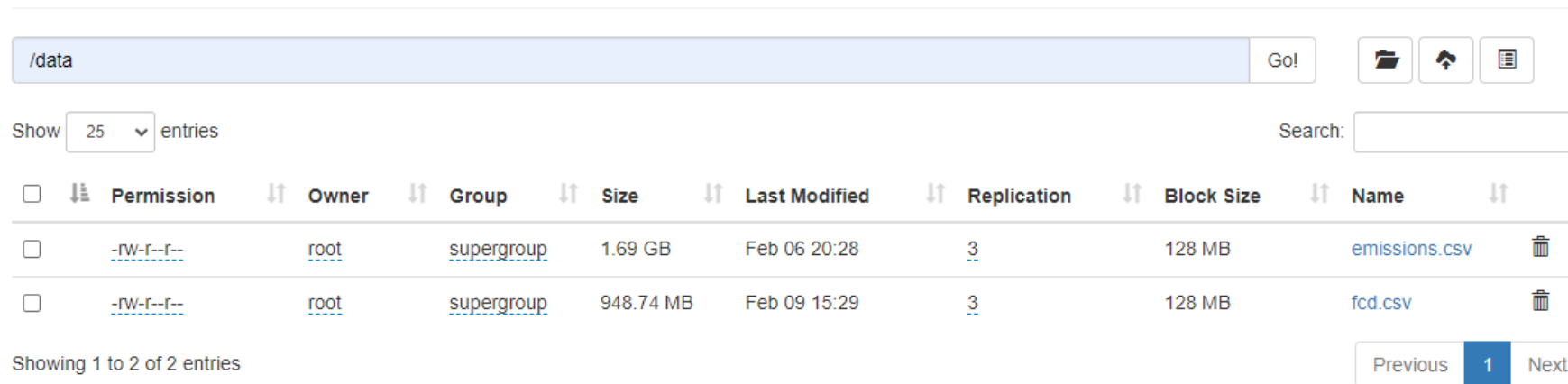
DOCKER COMPOSE

- Kroz docker se pokreću Hadoop kontejneri, kao i Spark



<input type="checkbox"/>	▼  projekat1		Running (8/8)	210.97%	12 seconds ago	  
<input type="checkbox"/>	 datanode e84996e4c28a 	bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8	Running	0.14%	13 seconds ago	  
<input type="checkbox"/>	 historyserver 565b6b516a56 	bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8	Running	0.91%	13 seconds ago	  
<input type="checkbox"/>	 namenode 22ff2c3918eb 	bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8	Running	0.13% 9000:9000  Show all ports (2)	13 seconds ago	  
<input type="checkbox"/>	 nodemanager c9b77b847c85 	bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8	Running	0%	13 seconds ago	  
<input type="checkbox"/>	 resourcemanager 96984817c10c 	bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8	Running	209.61%	13 seconds ago	  
<input type="checkbox"/>	 spark-master 8b1c3f6b5cc1 	bde2020/spark-master:3.1.2-hadoop3.2	Running	0.06% 7077:7077  Show all ports (2)	13 seconds ago	  
<input type="checkbox"/>	 spark-worker-1 29171aab2115 	bde2020/spark-worker:3.1.2-hadoop3.2	Running	0.06% 8071:8071 	12 seconds ago	  
<input type="checkbox"/>	 spark-worker-2 e4b818e91940 	bde2020/spark-worker:3.1.2-hadoop3.2	Running	0.06% 8072:8071 	13 seconds ago	  

POSTAVLJANJE PODATAKA NA HDFS

- Prvo na namenode:
 - `docker cp emissions.csv namenode:/data`
- A zatim i na HDFS:
 - `docker exec -it namenode bash`
 - `root@22ff2c3918eb:/# hdfs dfs -mkdir /data`
 - `root@22ff2c3918eb:/# hdfs dfs -put /data/emission.csv /data`



The screenshot shows the HDFS web interface for the /data directory. At the top, there is a search bar with the text "/data" and a "Go!" button. Below the search bar, there are three icons: a folder, an upload arrow, and a document. The main content area displays a table of files. The table has columns for checkboxes, Permission, Owner, Group, Size, Last Modified, Replication, Block Size, Name, and a trash icon. Two files are listed: "emissions.csv" and "fcd.csv". The "emissions.csv" file is 1.69 GB, created on Feb 06 20:28, with 3 replications. The "fcd.csv" file is 948.74 MB, created on Feb 09 15:29, with 3 replications. At the bottom left, it says "Showing 1 to 2 of 2 entries". At the bottom right, there are navigation buttons: "Previous", "1" (highlighted), and "Next".

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	-rw-r--r--	root	supergroup	1.69 GB	Feb 06 20:28	3	128 MB	emissions.csv	
<input type="checkbox"/>	-rw-r--r--	root	supergroup	948.74 MB	Feb 09 15:29	3	128 MB	fcd.csv	

Showing 1 to 2 of 2 entries

Previous 1 Next

APLIKACIJA

TASK 1

```
def find_vehicles(spark, latitude_point, longitude_point, proximity_size, time_start, time_end):  
  
    start_time = time.time()  
    data = spark.read.option("header", "true").option("delimiter", ",").csv('hdfs://namenode:9000/data/fcd.csv', inferSchema=True)  
  
    data_filter = data.filter((col("timestep_time") >= lit(time_start)) & (col("timestep_time") <= lit(time_end)))  
  
    target_latitude = radians(lit(latitude_point))  
    target_longitude = radians(lit(longitude_point))  
  
    df_with_distance = data_filter.withColumn(  
        "distance",  
        6371 * acos(  
            sin(target_latitude) * sin(radians(col("vehicle_x"))) +  
            cos(target_latitude) * cos(radians(col("vehicle_x"))) *  
            cos(radians(col("vehicle_y")) - target_longitude)  
        )  
    )  
  
    df_filtered = df_with_distance.filter(col("distance") < proximity_size).dropDuplicates(["vehicle_id"])  
    print(f"Broj jedinstvenih vozila u okolini: {df_filtered.count()}")  
  
    df_filtered.show()  
  
    return time.time() - start_time
```

Za računanje udaljenosti između
2 tačke korišćena je Haversine
formule

TASK 2

```
def air_fuel_task(spark, task, time_start, time_end):  
  
    start_time = time.time()  
  
    data = spark.read.option("header", "true").option("delimiter", ";").csv("hdfs://namenode:9000/data/emissions.csv", inferSchema=True)  
  
    data_filter = data.filter((col("timestep_time") >= time_start) & (col("timestep_time") <= time_end))  
  
    columns = []  
  
    if task == 'z':  
        # zagadjenje  
        columns = ["vehicle_CO", "vehicle_CO2", "vehicle_HC", "vehicle_NOx", "vehicle_PMX", "vehicle_noise"]  
    elif task == 'g':  
        # gorivo  
        columns = ["vehicle_electricity", "vehicle_fuel"]  
  
    for col_name in columns:  
        data_filter.groupBy("vehicle_lane") \  
            .agg(  
                count(col(col_name)).alias(f"count_{col_name}"), \  
                min(col(col_name)).alias(f"min_{col_name}"), \  
                max(col(col_name)).alias(f"max_{col_name}"), \  
                mean(col(col_name)).alias(f"mean_{col_name}"), \  
                avg(col(col_name)).alias(f"avg_{col_name}"), \  
                stddev(col(col_name)).alias(f"sddev_{col_name}") \  
            ) \  
            .show()  
  
    return time.time() - start_time
```

MAIN

```
if __name__ == "__main__":  
    ts = time.time()  
    #spark = SparkSession.builder.appName("App").master("spark://8b1c3f6b5cc1:7077").getOrCreate()  
    spark = SparkSession.builder.appName("AppLocal").master("local[2]").getOrCreate()  
    #spark = SparkSession.builder.appName("App").getOrCreate()  
  
    args = sys.argv  
    print(args)  
    t=0  
  
    if args[1] == '1':  
        print("task 1")  
        t = find_vehicles(spark,float(args[3]),float(args[2]),float(args[4]), float(args[5]), float(args[6]))  
  
    elif args[1] == '2':  
        print("task 2")  
        t = air_fuel_task(spark, args[2],float(args[3]),float(args[4]))  
  
    print(f"Vreme izvršenja task{args[1]}: {t}")  
    print(f"Vreme izvršenja aplikacije: {time.time()-ts}")  
    spark.stop()
```

Na klasteru

Lokalno

POKRETANJE APLIKACIJE

DOCKERFILE

 Dockerfile

```
1 FROM bde2020/spark-python-template:3.1.2-hadoop3.2
2
3 ENV SPARK_APPLICATION_PYTHON_LOCATION /app/main.py
4 ENV SPARK_APPLICATION_ARGS "2 z 0 500"
```

KREIRANJE DOCKER IMAGE

```
D:\BigData\Projekat1>docker build --rm -t bde/spark-app .
[+] Building 9.9s (9/9) FINISHED                                docker:default
=> [internal] load build definition from Dockerfile             0.0s
=> => transferring dockerfile: 189B                             0.0s
=> [internal] load .dockerignore                               0.0s
=> => transferring context: 2B                                   0.0s
=> [internal] load metadata for docker.io/bde2020/spark-python-template:3.1.2-hadoop3.2 1.4s
=> [1/1] FROM docker.io/bde2020/spark-python-template:3.1.2-hadoop3.2@sha256:905fcb3cf7890faaa21789aefb2e330d921 0.0s
=> [internal] load build context                               0.6s
=> => transferring context: 873.56kB                             0.6s
=> CACHED [2/1] COPY requirements.txt /app/                    0.0s
=> CACHED [3/1] RUN cd /app      && pip3 install -r requirements.txt 0.0s
=> [4/1] COPY . /app                                           5.1s
=> exporting to image                                           2.7s
=> => exporting layers                                           2.7s
=> => writing image sha256:16ffb8d280b481b16b33ab39b48dd60b51e19486cdc808d8dea71607280faa68 0.0s
=> => naming to docker.io/bde/spark-app                        0.0s
```

What's Next?

View summary of image vulnerabilities and recommendations → [docker scout quickview](#)

POKRETANJE APLIKACIJE

Postavljanje argumenata

```
D:\BigData\Projekat1>set SPARK_APPLICATION_ARGS=2 z 0 500
```

Pokretanje aplikacije

```
D:\BigData\Projekat1>docker run --rm --net bde -e SPARK_APPLICATION_ARGS -p 4040:4040 bde/spark-app
```

```
Submit application /app/main.py to Spark master spark://spark-master:7077
```

```
Passing arguments 2 z 0 500
```

```
24/02/09 16:57:10 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin
```

```
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
```

```
24/02/09 16:57:11 INFO SparkContext: Running Spark version 3.1.2
```

```
24/02/09 16:57:11 INFO ResourceUtils: =====
```

```
24/02/09 16:57:11 INFO ResourceUtils: No custom resources configured for spark.driver.
```

```
24/02/09 16:57:11 INFO ResourceUtils: =====
```

```
24/02/09 16:57:11 INFO SparkContext: Submitted application: App
```

```
24/02/09 16:57:11 INFO ResourceProfile: Default ResourceProfile created, executor resources: Map(cores -> n
```

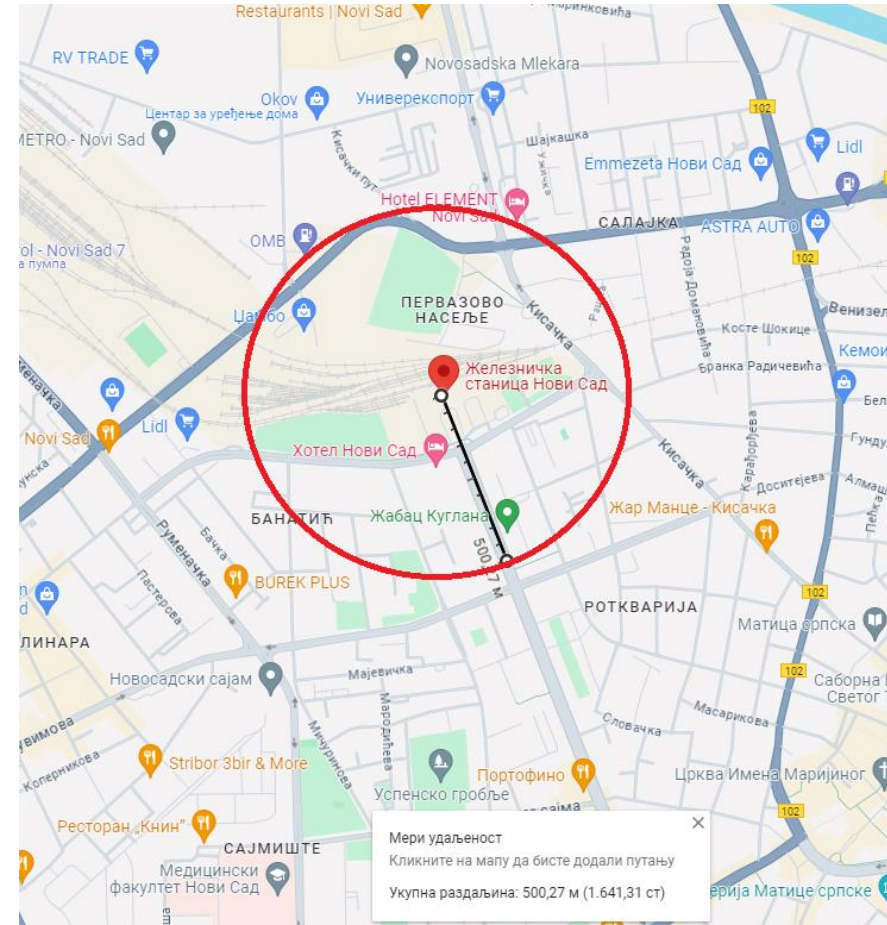

REZULTATI

TASK 1 - KLASITER

Pronazi sva vozila koja su prošla pored Železničke stanice u Novom Sadu, radius 500m, u vremenskom intervalu 0-500.

```
D:\BigData\Projekat1>set SPARK_APPLICATION_ARGS=1 45.26560669685613 19.829547417546962 0.5 0 500
```

```
24/02/00 16:54:33 INFO SparkContext: 888  
Broj jedinstvenih vozila u okolini: 92  
24/02/00 16:54:33 INFO FileSourceStrategy:
```



TASK 1 - KLAUSTER

24/02/09 16:54:38 INFO CodeGenerator: Code generated in 54.057498 ms

timestep_time	vehicle_angle	vehicle_id	vehicle_lane	vehicle_pos	vehicle_slope	vehicle_speed	vehicle_type	vehicle_x	vehicle_y	distance
362.0	90.84	veh421	8103927#3_1	186.72	0.0	14.04	veh_passenger	19.82768	45.269938	0.498382221850141
20.0	255.75	veh61	682003435#0_2	0.2	0.0	0.0	veh_passenger	19.82988	45.264049	0.16708175518650029
377.0	249.03	veh845	-272030397#0_0	125.78	0.0	14.46	veh_passenger	19.833473	45.26332	0.4977430150821242
80.0	265.09	veh119	8080984#0_2	24.14	0.0	13.43	veh_passenger	19.832287	45.269373	0.4979972561138862
192.0	265.09	veh157	8080984#0_2	27.0	0.0	10.89	veh_passenger	19.832251	45.269371	0.4953924340837514
287.0	265.09	veh789	8080984#0_2	25.81	0.0	12.94	veh_passenger	19.832266	45.269372	0.49648934448140497
356.0	90.84	veh504	8103927#3_1	193.35	0.0	15.29	veh_passenger	19.827764	45.269938	0.49456340978802715
132.0	245.75	veh420	725388592#2_1	1.14	0.0	0.0	veh_passenger	19.830644	45.26427	0.18552016330403515
175.0	265.09	veh274	8080984#0_2	27.37	0.0	14.47	veh_passenger	19.832246	45.26937	0.49497206851471537
102.0	165.96	veh132	245159912#10_0	69.71	0.0	13.75	veh_passenger	19.831451	45.269914	0.49779287064146394
473.0	95.18	veh1358	753847962#3_0	0.79	0.0	12.32	veh_passenger	19.825532	45.263746	0.4870724460015561
277.0	95.82	veh834	:cluster_38951754...	18.32	0.0	12.13	veh_passenger	19.826259	45.262442	0.4932431357822082
240.0	250.7	veh768	725388590_1	0.35	0.0	0.0	veh_passenger	19.83001	45.264103	0.16548594937543673
439.0	95.19	veh1257	:cluster_12296531...	6.05	0.0	14.52	veh_passenger	19.82552	45.263746	0.48829590990357896
231.0	177.24	veh544	:60493683_0_1	3.2	0.0	14.19	veh_passenger	19.831491	45.269885	0.49696697111805005
465.0	265.1	veh1305	8080984#0_3	19.84	0.0	13.18	veh_passenger	19.832346	45.269348	0.4999876864330848
175.0	240.23	veh508	682003446#2_0	76.68	0.0	13.5	veh_passenger	19.833896	45.265621	0.4835426307650172
447.0	255.75	veh1432	682003435#0_2	0.2	0.0	0.0	veh_passenger	19.82988	45.264049	0.16708175518650029
437.0	277.17	veh1398	32617696#2-Added0...	5.1	0.0	0.0	veh_passenger	19.827862	45.269997	0.49600396246211875
123.0	167.97	veh394	725805933#0_1	5.1	0.0	0.0	veh_passenger	19.831492	45.269793	0.4883684239486609

only showing top 20 rows

Vreme izvršenja task1: 28.32181406021118

Vreme izvršenja aplikacije: 30.895284414291382

24/02/09 16:54:38 INFO SparkUI: Stopped Spark web UI at http://46c0485e98de:4040

TASK 2 ZAGAĐENJE - KLASITER

vehicle_lane	count_vehicle_CO	min_vehicle_CO	max_vehicle_CO	mean_vehicle_CO	avg_vehicle_CO	sddev_vehicle_CO
:cluster_29864227...	4	0.0	0.0	0.0	0.0	0.0
-698946635#5_0	131	0.0	180.73	25.046183206106864	25.046183206106864	39.28809619010702
:3069940788_7_0	6	79.15	139.06	104.20833333333336	104.20833333333336	22.049189932209902
725331468_1	182	0.0	246.05	30.996098901098907	30.996098901098907	51.61721928471445
:2189608399_2_0	4	0.0	114.2	36.222500000000004	36.222500000000004	53.9605969666756
:9611216349_3_0	1	0.0	0.0	0.0	0.0	null
:cluster_16058699...	13	0.0	79.66	15.686153846153847	15.686153846153847	28.21252042931221
:904119318_7_0	9	0.0	56.01	8.19	8.19	18.55831484806743
8676539#0_0	7	0.0	15.39	4.132857142857143	4.132857142857143	7.07837249251953
:cluster_29842561...	5	0.0	64.87	33.346000000000004	33.346000000000004	27.869238956239908
:2285516135_3_1	27	0.0	147.99	57.09777777777777	57.09777777777777	57.75651087728508
:3069940788_3_0	4	0.0	33.92	8.48	8.48	16.96
-115158835#4_0	47	0.0	172.62	85.44297872340425	85.44297872340425	52.51318550425108
-144715839#2_0	24	0.0	185.22	67.71666666666668	67.71666666666668	61.65380616387568
:cluster_23207031...	15	0.0	146.65	25.942	25.942	50.056827591962204
1239699006_0	27	0.0	164.79	96.05555555555559	96.05555555555559	46.75515219107148
:cluster_11040490...	4	0.0	40.45	16.325000000000003	16.325000000000003	18.707229440334913
:cluster_10768597...	4	0.0	151.8	84.305	84.305	70.75453413032977
360741092#3_0	11	0.0	68.8	17.59181818181818	17.59181818181818	26.238705310354707
-105296041#5_0	19	0.0	163.61	22.09631578947368	22.09631578947368	44.65076709687282

only showing top 20 rows

TASK 2 ZAGAĐENJE - KLASITER

vehicle_lane	count_vehicle_CO2	min_vehicle_CO2	max_vehicle_CO2	mean_vehicle_CO2	avg_vehicle_CO2	sddev_vehicle_CO2
:cluster_29864227...	4	0.0	2806.07	701.5175	701.5175	1403.035
-698946635#5_0	131	0.0	11485.42	2581.643969465649	2581.643969465649	2260.7256624495526
:3069940788_7_0	6	3809.68	6298.9	4865.625	4865.625	970.3249731868185
725331468_1	182	0.0	15558.08	3604.462857142859	3604.462857142859	3409.7407443437983
:2189608399_2_0	4	0.0	6368.37	2502.1275	2502.1275	2729.2986331458246
:9611216349_3_0	1	0.0	0.0	0.0	0.0	null
:cluster_16058699...	13	0.0	6571.07	2258.1715384615386	2258.1715384615386	2300.350453513718
:904119318_7_0	9	0.0	5250.21	1563.0966666666666	1563.0966666666666	1841.2238694276155
8676539#0_0	7	0.0	3265.61	1615.987142857143	1615.987142857143	1319.322516340796
:cluster_29842561...	5	0.0	4509.8	2759.556	2759.556	1894.7885528285208
:2285516135_3_1	27	0.0	6523.96	2796.899259259259	2796.899259259259	2069.4101571535302
:3069940788_3_0	4	0.0	3525.22	1870.9074999999998	1870.9074999999998	1448.5406816833508
-115158835#4_0	47	0.0	10480.41	5698.172127659573	5698.172127659573	3016.885798898489
-144715839#2_0	24	0.0	11583.42	3973.56375	3973.56375	3638.716127995903
:cluster_23207031...	15	0.0	9361.58	3370.6320000000005	3370.6320000000005	2757.01080542004
1239699006_0	27	0.0	10155.06	5396.727407407408	5396.727407407408	2907.2965648474506
:cluster_11040490...	4	0.0	4742.86	3037.3199999999997	3037.3199999999997	2190.9478888067297
:cluster_10768597...	4	0.0	7995.49	4631.875	4631.875	3659.7535918373883
360741092#3_0	11	0.0	5722.4	2039.7590909090907	2039.7590909090907	2253.2865564834606
-105296041#5_0	19	0.0	9648.99	2426.8136842105264	2426.8136842105264	2744.059306959937

TASK 2 ZAGAĐENJE - KLASITER

24/02/09 16:57:51 INFO DAGScheduler: Job 4 finished: showString at NativeMethodAccessorImpl.java:0, took 4.104647 s

vehicle_lane	count_vehicle_HC	min_vehicle_HC	max_vehicle_HC	mean_vehicle_HC	avg_vehicle_HC	sddev_vehicle_HC
:cluster_29864227...	4	0.0	0.05	0.0125	0.0125	0.025
-698946635#5_0	131	0.0	1.1	0.16946564885496176	0.16946564885496176	0.22643835288456718
:3069940788_7_0	6	0.45	0.78	0.5883333333333334	0.5883333333333334	0.1215593133686871
725331468_1	182	0.0	1.5	0.2179670329670329	0.2179670329670329	0.3203826079576664
:2189608399_2_0	4	0.0	0.67	0.21500000000000002	0.21500000000000002	0.316280466253187
:9611216349_3_0	1	0.0	0.0	0.0	0.0	null
:cluster_16058699...	13	0.0	0.52	0.11923076923076925	0.11923076923076925	0.181496627078923
:904119318_7_0	9	0.0	0.38	0.06555555555555556	0.06555555555555556	0.12758439472669758
8676539#0_0	7	0.0	0.14	0.038571428571428576	0.038571428571428576	0.06593647736738024
:cluster_29842561...	5	0.0	0.4	0.21400000000000002	0.21400000000000002	0.1703819239238717
:2285516135_3_1	27	0.0	0.78	0.3225925925925926	0.3225925925925926	0.3048781924813781
:3069940788_3_0	4	0.0	0.24	0.065	0.065	0.11704699910719625
-115158835#4_0	47	0.0	1.04	0.5242553191489361	0.5242553191489361	0.30992106064252434
-144715839#2_0	24	0.0	1.12	0.4029166666666667	0.4029166666666667	0.3651501081052495
:cluster_23207031...	15	0.0	0.89	0.17866666666666667	0.17866666666666667	0.3008290132917622
1239699006_0	27	0.0	1.0	0.5655555555555556	0.5655555555555556	0.2794408336073658
:cluster_11040490...	4	0.0	0.3	0.1475	0.1475	0.13351029922818689
:cluster_10768597...	4	0.0	0.88	0.4925	0.4925	0.4080339038200952
360741092#3_0	11	0.0	0.45	0.12363636363636363	0.12363636363636363	0.17089603431751876
-105296041#5_0	19	0.0	0.98	0.1505263157894737	0.1505263157894737	0.2684414623590221

only showing top 20 rows

TASK 2 ZAGAĐENJE - KLASITER

vehicle_lane	count_vehicle_PmX	min_vehicle_PmX	max_vehicle_PmX	mean_vehicle_PmX	avg_vehicle_PmX	sddev_vehicle_PmX
:cluster_29864227...	4	0.0	0.02	0.005	0.005	0.01
-698946635#5_0	131	0.0	0.25	0.03541984732824425	0.03541984732824425	0.04463974672496272
:3069940788_7_0	6	0.06	0.13	0.09000000000000001	0.09000000000000001	0.026076809620810597
725331468_1	182	0.0	0.36	0.05747252747252743	0.05747252747252743	0.07627506788867447
:2189608399_2_0	4	0.0	0.13	0.0375	0.0375	0.062383224240709675
:9611216349_3_0	1	0.0	0.0	0.0	0.0	null
:cluster_16058699...	13	0.0	0.12	0.028461538461538465	0.028461538461538465	0.041602514716892185
:904119318_7_0	9	0.0	0.09	0.015555555555555553	0.015555555555555553	0.03004626062886658
8676539#0_0	7	0.0	0.04	0.01	0.01	0.017320508075688773
:cluster_29842561...	5	0.0	0.07	0.036	0.036	0.03361547262794322
:2285516135_3_1	27	0.0	0.13	0.04666666666666667	0.04666666666666667	0.04402796314232057
:3069940788_3_0	4	0.0	0.05	0.015000000000000001	0.015000000000000001	0.02380476142847617
-115158835#4_0	47	0.0	0.23	0.10936170212765957	0.10936170212765957	0.0674506020638727
-144715839#2_0	24	0.0	0.25	0.0775	0.0775	0.07605775837280279
:cluster_23207031...	15	0.0	0.2	0.04733333333333334	0.04733333333333334	0.06507870692659846
1239699006_0	27	0.0	0.22	0.10444444444444445	0.10444444444444445	0.0647678299379545
:cluster_11040490...	4	0.0	0.07	0.0425	0.0425	0.03403429642777023
:cluster_10768597...	4	0.0	0.17	0.0925	0.0925	0.08098353742170895
360741092#3_0	11	0.0	0.1	0.02727272727272727	0.02727272727272727	0.03690282078402926
-105296041#5_0	19	0.0	0.21	0.034736842105263156	0.034736842105263156	0.05824931292419122

only showing top 20 rows

TASK 2 ZAGAĐENJE - KLASITER

24/02/09 18:37:58 INFO DAGScheduler: Job 5 finished: showString at NativeMethodAccessorImpl.java:0, took 4.983079 s

vehicle_lane	count_vehicle_NOx	min_vehicle_NOx	max_vehicle_NOx	mean_vehicle_NOx	avg_vehicle_NOx	sddev_vehicle_NOx
:cluster_29864227...	4	0.0	0.9	0.225	0.225	0.45
-698946635#5_0	131	0.0	5.01	0.9637404580152675	0.9637404580152675	0.9453622308429738
:3069940788_7_0	6	1.56	2.75	2.0633333333333335	2.0633333333333335	0.45802474460084214
725331468_1	182	0.0	6.82	1.3685714285714285	1.3685714285714285	1.4770581269331906
:2189608399_2_0	4	0.0	2.72	0.9525000000000001	0.9525000000000001	1.216480579376424
:9611216349_3_0	1	0.0	0.0	0.0	0.0	null
:cluster_16058699...	13	0.0	2.69	0.8246153846153846	0.8246153846153846	0.9274841902032203
:904119318_7_0	9	0.0	2.09	0.5277777777777778	0.5277777777777778	0.7131405502735378
8676539#0_0	7	0.0	1.15	0.4857142857142857	0.4857142857142857	0.4740554518498563
:cluster_29842561...	5	0.0	1.82	1.05	1.05	0.7777531742140305
:2285516135_3_1	27	0.0	2.81	1.1303703703703705	1.1303703703703705	0.8999037618694461
:3069940788_3_0	4	0.0	1.32	0.6074999999999999	0.6074999999999999	0.5435301279598032
-115158835#4_0	47	0.0	4.59	2.3870212765957453	2.3870212765957453	1.3313133673730468
-144715839#2_0	24	0.0	5.07	1.6854166666666668	1.6854166666666668	1.57246684949274
:cluster_23207031...	15	0.0	4.05	1.2093333333333334	1.2093333333333334	1.2556072330986008
1239699006_0	27	0.0	4.43	2.29037037037037	2.29037037037037	1.2780046992855743
:cluster_11040490...	4	0.0	1.83	1.105	1.105	0.8280297901243249
:cluster_10768597...	4	0.0	3.5	1.9775	1.9775	1.615392109262227
360741092#3_0	11	0.0	2.32	0.7554545454545454	0.7554545454545454	0.8979461413287952
-105296041#5_0	19	0.0	4.22	0.9021052631578947	0.9021052631578947	1.1744766160453886

only showing top 20 rows

TASK 2 ZAGADENJE - KLASITER

vehicle_lane	count_vehicle_noise	min_vehicle_noise	max_vehicle_noise	mean_vehicle_noise	avg_vehicle_noise	sddev_vehicle_noise
:cluster_29864227...	4	60.9	65.51	62.8475	62.8475	1.9478428239123087
-698946635#5_0	131	55.94	73.1	64.1590076335878	64.1590076335878	2.506703276282591
:3069940788_7_0	6	63.28	69.11	65.59833333333333	65.59833333333333	2.16893906476569
725331468_1	182	61.63	75.74	66.69857142857146	66.69857142857146	2.6114744978423237
:2189608399_2_0	4	59.86	67.8	62.897499999999994	62.897499999999994	3.4981554663374603
:9611216349_3_0	1	60.92	60.92	60.92	60.92	null
:cluster_16058699...	13	61.83	68.03	64.49538461538461	64.49538461538461	1.900546480114843
:904119318_7_0	9	61.17	66.66	63.224444444444444	63.224444444444444	1.8113883564210558
8676539#0_0	7	54.29	64.99	62.30000000000001	62.30000000000001	3.785573492440298
:cluster_29842561...	5	59.64	65.03	62.878	62.878	2.316456345368934
:2285516135_3_1	27	54.91	68.31	63.39888888888888	63.39888888888888	3.0670950091041838
:3069940788_3_0	4	61.87	64.26	63.3775	63.3775	1.1363501514351388
-115158835#4_0	47	55.98	72.65	67.11595744680851	67.11595744680851	3.396612544986336
-144715839#2_0	24	37.48	73.35	61.62583333333331	61.62583333333331	9.138786864913683
:cluster_23207031...	15	63.1	71.07	66.52333333333333	66.52333333333333	2.226697319435436
1239699006_0	27	56.4	72.16	66.2651851851852	66.2651851851852	4.170098108392438
:cluster_11040490...	4	63.38	68.01	65.75999999999999	65.75999999999999	1.9821368940279254
:cluster_10768597...	4	59.2	70.93	65.63	65.63	5.431261363624477
360741092#3_0	11	61.44	66.98	63.94727272727275	63.94727272727275	1.8068264493807442
-105296041#5_0	19	61.74	72.04	64.62421052631579	64.62421052631579	2.5631259469411645

only showing top 20 rows

Vreme izvršenja task2: 51.10928153991699

Vreme izvršenja aplikacije: 54.176167726516724

24/02/09 16:58:05 INFO SparkUI: Stopped Spark web UI at http://8f8d15a3ac05:4040

24/02/09 16:58:05 INFO StandaloneSchedulerBackend: Shutting down all executors

24/02/09 16:58:05 INFO CoarseGrainedSchedulerBackend\$DriverEndpoint: Asking each executor to shut down

TASK 2 GORIVO - KLAUSTER

```
24/02/09 16:59:12 INFO DAGScheduler: Job 3 finished: showString at NativeMethodAccessorImpl.java:0, took 4.795363 s
```

vehicle_lane	count_vehicle_fuel	min_vehicle_fuel	max_vehicle_fuel	mean_vehicle_fuel	avg_vehicle_fuel	sddev_vehicle_fuel
:cluster_29864227...	4	0.0	895.02	223.755	223.755	447.51
-698946635#5_0	131	0.0	3663.27	823.4293893129767	823.4293893129767	721.06274816013
:3069940788_7_0	6	1215.13	2009.06	1551.9199999999998	1551.9199999999998	309.4826538596307
725331468_1	182	0.0	4962.23	1149.6528021978024	1149.6528021978024	1087.5345370596874
:2189608399_2_0	4	0.0	2031.21	798.0725	798.0725	870.5136172542047
:9611216349_3_0	1	0.0	0.0	0.0	0.0	null
:cluster_16058699...	13	0.0	2095.85	720.2553846153846	720.2553846153846	733.7019833308273
:904119318_7_0	9	0.0	1674.57	498.56111111111113	498.56111111111113	587.26580488405
8676539#0_0	7	0.0	1041.59	515.4342857142857	515.4342857142857	420.8063516574317
:cluster_29842561...	5	0.0	1438.42	880.18	880.18	604.3500651526399
:2285516135_3_1	27	0.0	2080.84	892.094074074074	892.094074074074	660.0485203625286
:3069940788_3_0	4	0.0	1124.39	596.7425000000001	596.7425000000001	462.02093057746504
-115158835#4_0	47	0.0	3342.73	1817.4440425531914	1817.4440425531914	962.234011867265
-144715839#2_0	24	0.0	3694.53	1267.3795833333336	1267.3795833333336	1160.5718303022538
:cluster_23207031...	15	0.0	2985.88	1075.0753333333332	1075.0753333333332	879.3466471101199
1239699006_0	27	0.0	3238.96	1721.3022222222223	1721.3022222222223	927.2787487022831
:cluster_11040490...	4	0.0	1512.75	968.7649999999999	968.7649999999999	698.809521996755
:cluster_10768597...	4	0.0	2550.18	1477.3525	1477.3525	1167.2865937542217
360741092#3_0	11	0.0	1825.17	650.59	650.59	718.691543375877
-105296041#5_0	19	0.0	3077.55	774.0436842105262	774.0436842105262	875.2201632491649

```
only showing top 20 rows
```

Vreme izvršenja task2: 29.273178100585938
Vreme izvršenja aplikacije: 32.55281591415405

```
24/02/09 16:59:12 INFO SparkUI: Stopped Spark web UI at http://cb912081b08a:4040
```

SPARK MASTER



Spark Master at spark://8b1c3f6b5cc1:7077

URL: spark://8b1c3f6b5cc1:7077

Alive Workers: 2

Cores in use: 12 Total, 0 Used

Memory in use: 13.5 GiB Total, 0.0 B Used

Resources in use:

Applications: 0 Running, 5 Completed

Drivers: 0 Running, 0 Completed

Status: ALIVE

Task 2 gorivo

Task 2 zagađenje

Task 1

Workers (2)

Worker Id	Address	State	Cores	Memory	Resources
worker-20240209200018-172.19.0.8-43941	172.19.0.8:43941	ALIVE	6 (0 Used)	6.7 GiB (0.0 B Used)	
worker-20240209200018-172.19.0.9-38895	172.19.0.9:38895	ALIVE	6 (0 Used)	6.7 GiB (0.0 B Used)	

Running Applications (0)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
----------------	------	-------	---------------------	------------------------	----------------	------	-------	----------

Completed Applications (5)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
app-20240209201117-0004	App	12	1024.0 MiB		2024/02/09 20:11:17	root	FINISHED	36 s
app-20240209200630-0001	App	12	1024.0 MiB		2024/02/09 20:06:30	root	FINISHED	59 s
app-20240209200453-0000	App	12	1024.0 MiB		2024/02/09 20:04:53	root	FINISHED	34 s

TASK 1 - LOKALNO

timestep_time	vehicle_angle	vehicle_id	vehicle_lane	vehicle_pos	vehicle_slope	vehicle_speed	vehicle_type	vehicle_x	vehicle_y	distance
362.0	90.84	veh421	8103927#3_1	186.72	0.0	14.04	veh_passenger	19.82768	45.269938	0.49315134523133936
20.0	255.75	veh61	682003435#0_2	0.2	0.0	0.0	veh_passenger	19.82988	45.264049	0.17257615674884727
378.0	249.03	veh845	-272030397#0_0	140.71	0.0	14.92	veh_passenger	19.833296	45.26327	0.48647421028125654
80.0	265.09	veh119	8080984#0_2	24.14	0.0	13.43	veh_passenger	19.832287	45.269373	0.49398329579561523
192.0	265.09	veh157	8080984#0_2	27.0	0.0	10.89	veh_passenger	19.832251	45.269371	0.4913551407367016
287.0	265.09	veh789	8080984#0_2	25.81	0.0	12.94	veh_passenger	19.832266	45.269372	0.49246169360365954
355.0	87.13	veh504	8103927#3_1	178.06	0.0	15.21	veh_passenger	19.827569	45.269938	0.4984608626711837
132.0	245.75	veh420	725388592#2_1	1.14	0.0	0.0	veh_passenger	19.830644	45.26427	0.190055155395157
175.0	265.09	veh274	8080984#0_2	27.37	0.0	14.47	veh_passenger	19.832246	45.26937	0.49093188731103315
102.0	165.96	veh132	245159912#10_0	69.71	0.0	13.75	veh_passenger	19.831451	45.269914	0.49304350651431844
473.0	95.18	veh1358	753847962#3_0	0.79	0.0	12.32	veh_passenger	19.825532	45.263746	0.48880756920268226
277.0	95.82	veh834	:cluster_38951754...	18.32	0.0	12.13	veh_passenger	19.826259	45.262442	0.49656077999611786
240.0	250.7	veh768	725388590_1	0.35	0.0	0.0	veh_passenger	19.83001	45.264103	0.17089527099868265
439.0	95.19	veh1257	:cluster_12296531...	6.05	0.0	14.52	veh_passenger	19.82552	45.263746	0.49002524799867286
231.0	177.24	veh544	:60493683_0_1	3.2	0.0	14.19	veh_passenger	19.831491	45.269885	0.4922484825458502
465.0	265.1	veh1305	8080984#0_3	19.84	0.0	13.18	veh_passenger	19.832346	45.269348	0.4960260302168405
175.0	240.23	veh508	682003446#2_0	76.68	0.0	13.5	veh_passenger	19.833896	45.265621	0.4840970909113959
447.0	255.75	veh1432	682003435#0_2	0.2	0.0	0.0	veh_passenger	19.82988	45.264049	0.17257615674884727
437.0	277.17	veh1398	32617696#2-Added0...	5.1	0.0	0.0	veh_passenger	19.827862	45.269997	0.4907007101414445
123.0	167.97	veh394	725805933#0_1	5.1	0.0	0.0	veh_passenger	19.831492	45.269793	0.4836758061243394

only showing top 20 rows

Vreme izvršenja task1: 37.07720923423767

Vreme izvršenja aplikacije: 39.20816397666931

TASK 2 ZAGAĐENJE - LOKALNO

vehicle_lane	count_vehicle_noise	min_vehicle_noise	max_vehicle_noise	mean_vehicle_noise	avg_vehicle_noise	sddev_vehicle_noise
:cluster_29864227...	4	60.9	65.51	62.8475	62.8475	1.9478428239123087
-698946635#5_0	131	55.94	73.1	64.1590076335878	64.1590076335878	2.506703276282591
:3069940788_7_0	6	63.28	69.11	65.59833333333333	65.59833333333333	2.16893906476569
725331468_1	182	61.63	75.74	66.69857142857146	66.69857142857146	2.6114744978423237
:2189608399_2_0	4	59.86	67.8	62.897499999999994	62.897499999999994	3.4981554663374603
:9611216349_3_0	1	60.92	60.92	60.92	60.92	null
:cluster_16058699...	13	61.83	68.03	64.49538461538461	64.49538461538461	1.900546480114843
:904119318_7_0	9	61.17	66.66	63.224444444444444	63.224444444444444	1.8113883564210558
8676539#0_0	7	54.29	64.99	62.300000000000001	62.300000000000001	3.785573492440298
:cluster_29842561...	5	59.64	65.03	62.878	62.878	2.316456345368934
:2285516135_3_1	27	54.91	68.31	63.39888888888888	63.39888888888888	3.0670950091041838
:3069940788_3_0	4	61.87	64.26	63.3775	63.3775	1.1363501514351388
-115158835#4_0	47	55.98	72.65	67.11595744680851	67.11595744680851	3.396612544986336
-144715839#2_0	24	37.48	73.35	61.625833333333331	61.625833333333331	9.138786864913683
:cluster_23207031...	15	63.1	71.07	66.52333333333333	66.52333333333333	2.226697319435436
1239699006_0	27	56.4	72.16	66.2651851851852	66.2651851851852	4.170098108392438
:cluster_11040490...	4	63.38	68.01	65.75999999999999	65.75999999999999	1.9821368940279254
:cluster_10768597...	4	59.2	70.93	65.63	65.63	5.431261363624477
360741092#3_0	11	61.44	66.98	63.94727272727275	63.94727272727275	1.8068264493807442
-105296041#5_0	19	61.74	72.04	64.62421052631579	64.62421052631579	2.5631259469411645

only showing top 20 rows

Vreme izvršenja task2: 92.00439310073853
Vreme izvršenja aplikacije: 94.11160516738892

2024/02/09 20:20:38 INFO SparkUI: Stopped Spark web UI at http://3be18dee3b55:4040

TASK 2 GORIVO - LOKALNO

vehicle_lane	count_vehicle_fuel	min_vehicle_fuel	max_vehicle_fuel	mean_vehicle_fuel	avg_vehicle_fuel	sddev_vehicle_fuel
:cluster_29864227...	4	0.0	895.02	223.755	223.755	447.51
-698946635#5_0	131	0.0	3663.27	823.4293893129767	823.4293893129767	721.06274816013
:3069940788_7_0	6	1215.13	2009.06	1551.9199999999998	1551.9199999999998	309.4826538596307
725331468_1	182	0.0	4962.23	1149.6528021978024	1149.6528021978024	1087.5345370596874
:2189608399_2_0	4	0.0	2031.21	798.0725	798.0725	870.5136172542047
:9611216349_3_0	1	0.0	0.0	0.0	0.0	null
:cluster_16058699...	13	0.0	2095.85	720.2553846153846	720.2553846153846	733.7019833308273
:904119318_7_0	9	0.0	1674.57	498.56111111111113	498.56111111111113	587.26580488405
8676539#0_0	7	0.0	1041.59	515.4342857142857	515.4342857142857	420.8063516574317
:cluster_29842561...	5	0.0	1438.42	880.18	880.18	604.3500651526399
:2285516135_3_1	27	0.0	2080.84	892.094074074074	892.094074074074	660.0485203625286
:3069940788_3_0	4	0.0	1124.39	596.7425000000001	596.7425000000001	462.02093057746504
-115158835#4_0	47	0.0	3342.73	1817.4440425531914	1817.4440425531914	962.234011867265
-144715839#2_0	24	0.0	3694.53	1267.3795833333336	1267.3795833333336	1160.5718303022538
:cluster_23207031...	15	0.0	2985.88	1075.0753333333332	1075.0753333333332	879.3466471101199
1239699006_0	27	0.0	3238.96	1721.3022222222223	1721.3022222222223	927.2787487022831
:cluster_11040490...	4	0.0	1512.75	968.7649999999999	968.7649999999999	698.809521996755
:cluster_10768597...	4	0.0	2550.18	1477.3525	1477.3525	1167.2865937542217
360741092#3_0	11	0.0	1825.17	650.59	650.59	718.691543375877
-105296041#5_0	19	0.0	3077.55	774.0436842105262	774.0436842105262	875.2201632491649

only showing top 20 rows

Vreme izvršenja task2: 52.35433340072632

Vreme izvršenja aplikacije: 55.10133385658264

24/02/09 20:17:34 INFO BlockManagerInfo: Removed broadcast_9_piece0 on 86ffb20af7d4:32947 in memory (size: 22.7 KiB, free: 366.2 MiB)

POREĐENJE REZULTATA

	KLASTER	LOKALNO
Task 1	30,89s	39,20s
Task 2 Zagađenje	54,17s	94,11s
Task 2 Gorivo	32,55s	55,10s