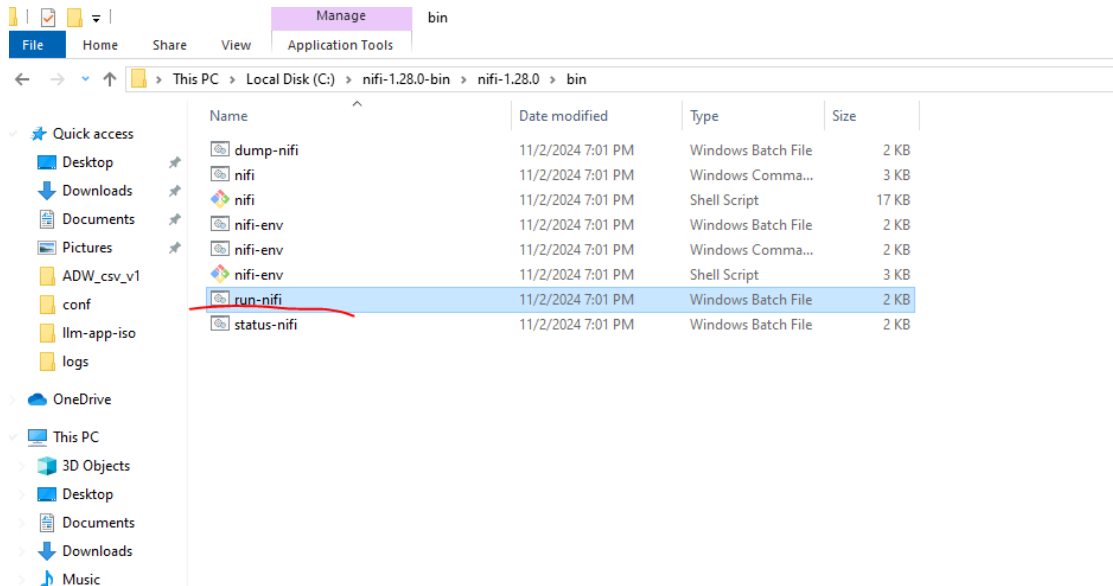


BDA Big DATA Infrastructure - Khawaja Abdullah

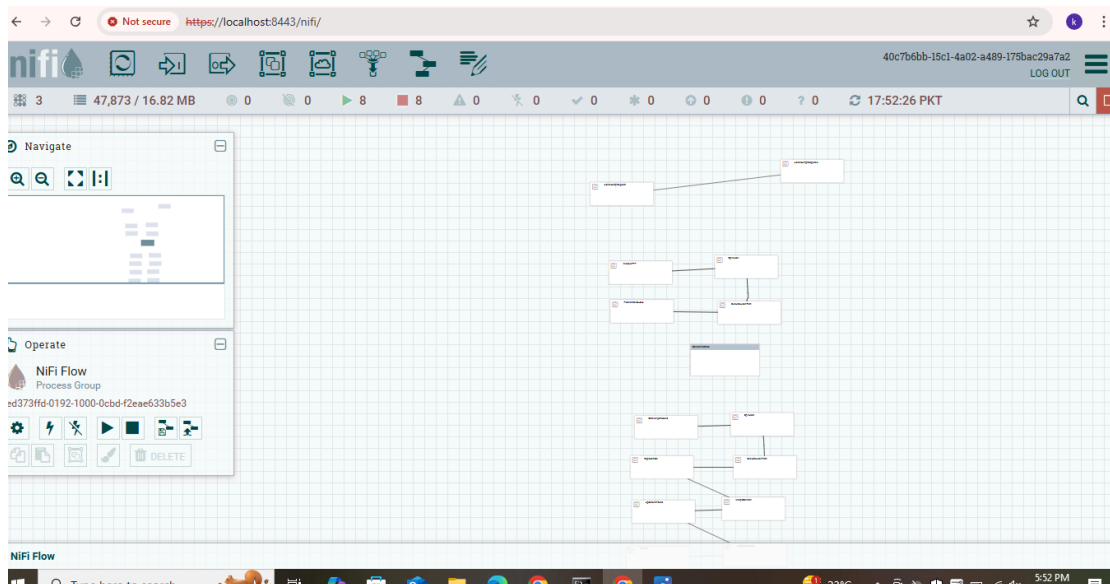
NIFI SETUP

Run NIFI from here:



```
C:\WINDOWS\system32\cmd.exe
2025-01-21 17:38:59,472 INFO [main] org.apache.nifi.bootstrap.Command Starting Apache NiFi...
2025-01-21 17:38:59,473 INFO [main] org.apache.nifi.bootstrap.Command Working Directory: C:\NIFI-1~1.0-B\NIFI-1~1.0
2025-01-21 17:38:59,479 INFO [main] org.apache.nifi.bootstrap.Command Command: C:\Program Files\Java\jdk-1.8\bin\java.exe
-e -classpath C:\NIFI-1~1.0-B\NIFI-1~1.0\conf;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\javax.servlet-api-3.1.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\jcl-over-slf4j-2.0.16.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\jetty-schemas-5.2.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\jul-to-slf4j-2.0.16.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\log4j-over-slf4j-2.0.16.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\logback-classic-1.3.14.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\logback-core-1.3.14.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-api-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-framework-api-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-nar-utils-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-per-process-group-logging-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-properties-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-property-utils-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-runtime-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-server-api-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-stateless-api-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\nifi-stateless-bootstrap-1.28.0.jar;C:\NIFI-1~1.0-B\NIFI-1~1.0\lib\slf4j-api-2.0.16.jar -Dorg.apache.jasper.compiler.disablejsr199=true -Xmx1g -Xms1g -Dcutorator-log-only-first-connection-issue-as-error-level=true -Djavax.security.auth.useSubjectCredsOnly=true -Djava.security.egd=file:/dev/urandom -Dzookeeper.admin.enableServer=false -Dsun.net.http.allowRestrictedHeaders=true -Djava.net.preferIPv4Stack=true -Djava.awt.headless=true -Djava.protocol.handler.pkgs=sun.net.www.protocol -Dnifi.properties.file.path=C:\NIFI-1~1.0-B\NIFI-1~1.0\conf\nifi.properties -Dnifi.bootstrap.listen.port=63775 -Dapp=NiFi -Dorg.apache.nifi.bootstrap.config.log.dir=C:\NIFI-1~1.0-B\NIFI-1~1.0\bin\..\logs org.apache.nifi.NiFi
2025-01-21 17:38:59,491 INFO [main] org.apache.nifi.bootstrap.Command Failed to determine Process ID from [java.lang.ProcessImpl]: java.lang.Process.pid()
2025-01-21 17:38:59,492 WARN [main] org.apache.nifi.bootstrap.Command Launched Apache NiFi but could not determined the Process ID
```

Login to NIFI:



SETUP KAFKA:

```
C:\Windows\System32\cmd.exe - .\bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties
Microsoft Windows [Version 10.0.19045.5247]
(c) Microsoft Corporation. All rights reserved.

C:\kafka>.bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties
[2025-01-21 17:57:41,067] INFO Reading configuration from: .\config\zookeeper.properties (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,117] INFO clientPortAddress is 0.0.0.0:2181 (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,118] INFO secureClientPort is not set (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,118] INFO observerMasterPort is not set (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,119] INFO metricsProvider.className is org.apache.zookeeper.metrics.impl.DefaultMetricsProvider (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,125] INFO autopurge.snapRetainCount set to 3 (org.apache.zookeeper.server.DataDirCleanupManager)
[2025-01-21 17:57:41,126] INFO autopurge.purgeInterval set to 0 (org.apache.zookeeper.server.DataDirCleanupManager)
[2025-01-21 17:57:41,126] INFO Purge task is not scheduled. (org.apache.zookeeper.server.DataDirCleanupManager)
[2025-01-21 17:57:41,126] WARN Either no config or no quorum defined in config, running in standalone mode (org.apache.zookeeper.server.quorum.QuorumPeerMain)
[2025-01-21 17:57:41,131] INFO Log4j 1.2 jmx support not found; jmx disabled. (org.apache.zookeeper.jmx.ManagedUtil)
[2025-01-21 17:57:41,134] INFO Reading configuration from: .\config\zookeeper.properties (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,137] INFO clientPortAddress is 0.0.0.0:2181 (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,137] INFO secureClientPort is not set (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,137] INFO observerMasterPort is not set (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,138] INFO metricsProvider.className is org.apache.zookeeper.metrics.impl.DefaultMetricsProvider (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2025-01-21 17:57:41,138] INFO Starting server (org.apache.zookeeper.server.ZooKeeperServerMain)
[2025-01-21 17:57:41,185] INFO ServerMetrics initialized with provider org.apache.zookeeper.metrics.impl.DefaultMetricsProvider@617c74e5 (org.apache.zookeeper.server.ServerMetrics)
[2025-01-21 17:57:41,216] INFO ACL digest algorithm is: SHA1 (org.apache.zookeeper.server.auth.DigestAuthenticationProvider)
```

```
C:\Windows\System32\cmd.exe - .\bin\windows\kafka-server-start.bat .\config\server.properties
Microsoft Windows [Version 10.0.19045.5247]
(c) Microsoft Corporation. All rights reserved.

C:\kafka>.\bin\windows\kafka-server-start.bat .\config\server.properties
[2025-01-21 17:59:12,523] INFO Registered kafka:type=kafka.Log4jController MBean (kafka.utils.Log4jControllerRegistration$)
[2025-01-21 17:59:13,573] INFO Setting -D jdk.tls.rejectClientInitiatedRenegotiation=true to disable client-initiated TLS renegotiation (org.apache.zookeeper.common.X509Util)
[2025-01-21 17:59:13,822] INFO starting (kafka.server.KafkaServer)
[2025-01-21 17:59:13,824] INFO Connecting to zookeeper on localhost:2181 (kafka.server.KafkaServer)
[2025-01-21 17:59:13,863] INFO [ZooKeeperClient Kafka server] Initializing a new session to localhost:2181. (kafka.zookeeper.ZooKeeperClient)
[2025-01-21 17:59:13,877] INFO Client environment:zookeeper.version=3.8.3-6ad6d364c7c0bcf0de452d54ebefa3058098ab56, built on 2023-10-05 10:34 UTC (org.apache.zookeeper.ZooKeeper)
[2025-01-21 17:59:13,878] INFO Client environment:host.name=DESKTOP-983IAU0.mshome.net (org.apache.zookeeper.ZooKeeper)
[2025-01-21 17:59:13,878] INFO Client environment:java.version=1.8.0_401 (org.apache.zookeeper.ZooKeeper)
[2025-01-21 17:59:13,878] INFO Client environment:java.vendor=Oracle Corporation (org.apache.zookeeper.ZooKeeper)
[2025-01-21 17:59:13,878] INFO Client environment:java.home=C:\Program Files\Java\jdk-1.8\jre (org.apache.zookeeper.ZooKeeper)
[2025-01-21 17:59:13,878] INFO Client environment:java.class.path=C:\kafka\libs\activation-1.1.1.jar;C:\kafka\libs\aaliance-repackaged-2.6.1.jar;C:\kafka\libs\argparse4j-0.7.0.jar;C:\kafka\libs\audience-annotations-0.12.0.jar;C:\kafka\libs\caffeine-2.9.3.jar;C:\kafka\libs\checker-qual-3.19.0.jar;C:\kafka\libs\commons-beanutils-1.9.4.jar;C:\kafka\libs\commons-cli-1.4.jar;C:\kafka\libs\commons-collections-3.2.2.jar;C:\kafka\libs\commons-digester-2.1.jar;C:\kafka\libs\commons-io-2.11.0.jar;C:\kafka\libs\commons-lang3-3.8.1.jar;C:\kafka\libs\commons-logging-1.2.jar;C:\kafka\libs\commons-validator-1.7.jar;C:\kafka\libs\connect-api-3.7.0.jar;C:\kafka\libs\connect-basic-auth-extension-3.7.0.jar;C:\kafka\libs\connect-file-3.7.0.jar;C:\kafka\libs\connect-json-3.7.0.jar;C:\kafka\libs\connect-mirror-3.7.0.jar;C:\kafka\libs\connect-mirror-client-3.7.0.jar;C:\kafka\libs\connect-runtime-3.7.0.jar;C:\kafka\libs\connect-transforms-3.7.0.jar;C:\kafka\libs\error-prone-annotations-2.10.0.jar;C:\kafka\libs\hk2-api-2.6.1.jar;C:\kafka\libs\hk2-locator-2.6.1.jar;C:\kafka\libs\hk2-utils-2.6.1.jar;C:\kafka\libs\jackson-annotations-2.16.0.jar;C:\kafka\libs\jackson-core-2.16.0.jar;C:\kafka\libs\jackson-data
```

Kafka Producer:

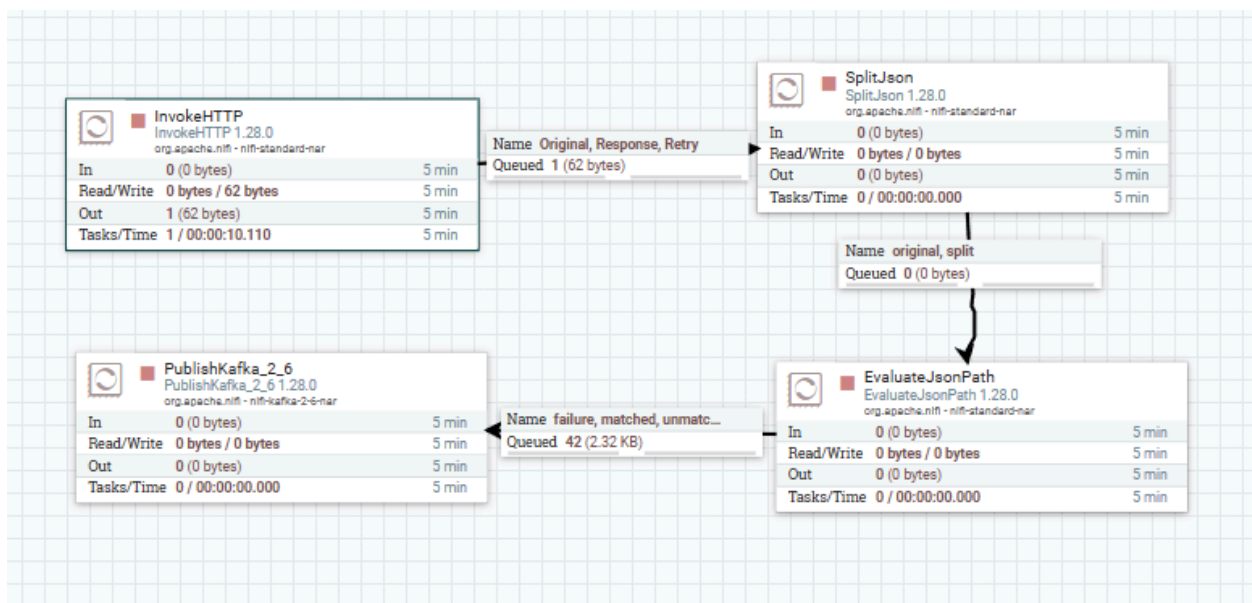
```
C:\Windows\System32\cmd.exe - .\bin\windows\kafka-console-producer.bat --topic medium-events --bootstrap-server localhost:9092
C:\kafka>.\bin\windows\kafka-console-producer.bat --topic medium-events --bootstrap-server localhost:9092
>
```

Kafka Consumer:

```
C:\kafka>.bin\windows\kafka-console-consumer.bat --topic medium-events --from-beginning --bootstrap-server localhost:9092
Hello
```

NIFI SETUP TO SET API DATA IN MONGODB

Add these Processors:



Invoke HTTP Processor:

Configure Processor | InvokeHTTP 1.28.0

Stopped

SETTINGS | SCHEDULING | PROPERTIES | RELATIONSHIPS | COMMENTS

Required field

Property	Value
HTTP Method	GET
HTTP URL	https://api.coinbase.com/v2/prices/btc-usd/spot
HTTP/2 Disabled	False
SSL Context Service	No value set
Socket Connect Timeout	5 secs
Socket Read Timeout	15 secs
Socket Write Timeout	15 secs
Socket Idle Timeout	5 mins
Socket Idle Connections	5
Proxy Configuration Service	No value set
Proxy Host	No value set
Request OAuth2 Access Token Provider	No value set

CANCEL APPLY

SplitJSON Processor:

Configure Processor | Split.Json 1.28.0

Stopped

SETTINGS | SCHEDULING | **PROPERTIES** | RELATIONSHIPS | COMMENTS

Required field ✔ +

Property	Value
JsonPath Expression	? \$.*
Null Value Representation	? empty string
Max String Length	? 20 MB

CANCEL APPLY

Evaluate JSON PATH Processor:

Configure Processor | Evaluate.JsonPath 1.28.0

Stopped

SETTINGS | SCHEDULING | **PROPERTIES** | RELATIONSHIPS | COMMENTS

Required field ✔ +

Property	Value
Destination	? flowfile-attribute
Return Type	? auto-detect
Path Not Found Behavior	? ignore
Null Value Representation	? empty string
Max String Length	? 20 MB
amount	? amount ✕
base	? base ✕
currency	? currency ✕

CANCEL APPLY

PublishKafka Processor

Configure Processor | PublishKafka_2_6.1.28.0

Stopped

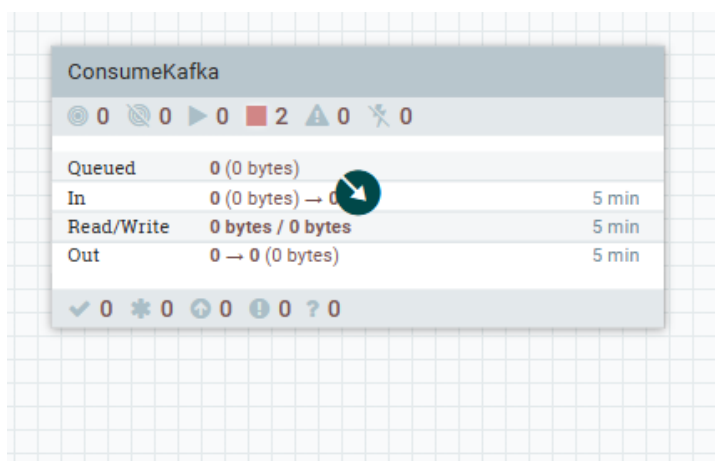
SETTINGS | SCHEDULING | **PROPERTIES** | RELATIONSHIPS | COMMENTS

Required field

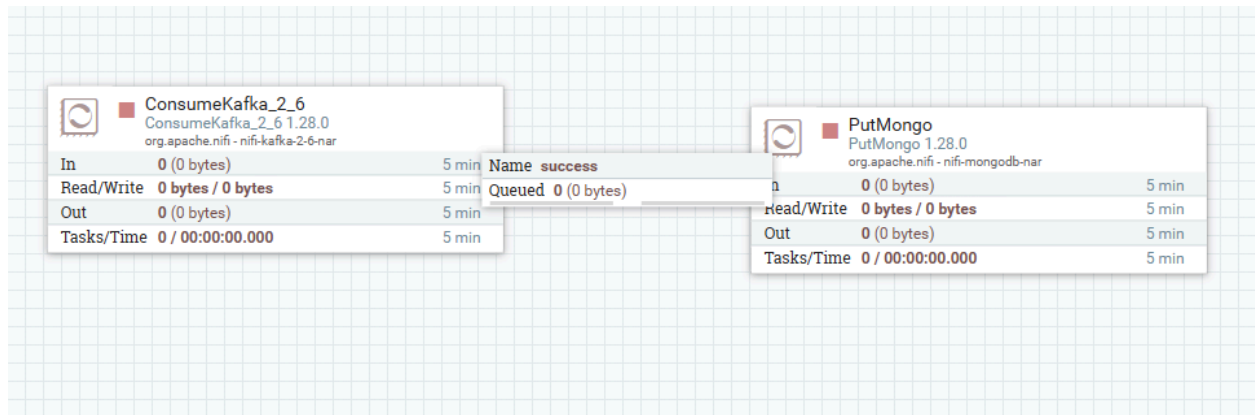
Property	Value
Kafka Brokers	localhost:9092
Topic Name	medium-events
Use Transactions	true
Transactional Id Prefix	No value set
Message Demarcator	No value set
Failure Strategy	Route to Failure
Delivery Guarantee	Guarantee Replicated Delivery
Attributes to Send as Headers (Regex)	No value set
Message Header Encoding	UTF-8
Security Protocol	PLAINTEXT
SASL Mechanism	GSSAPI
Kerberos Credentials Service	No value set

CANCELAPPLY

Then Add this Process Group:



Inside the Process group Add these Processors:



Consume Kafka Processor

Configure Processor | ConsumeKafka_2_6 1.28.0

Stopped

SETTINGS | SCHEDULING | **PROPERTIES** | RELATIONSHIPS | COMMENTS

Required field

Property	Value
Kafka Brokers	localhost:9092
Topic Name(s)	medium-events
Topic Name Format	names
Group ID	BTCRates
Commit Offsets	true
Max Uncommitted Time	1 secs
Honor Transactions	true
Message Demarcator	No value set
Separate By Key	false
Security Protocol	PLAINTEXT
SASL Mechanism	GSSAPI
Kerberos Credentials Service	No value set

CANCEL | APPLY

PutMongo Processor:

Configure Processor | PutMongo 1.28.0

Stopped

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Required field

Property

Value

Client Service	No value set
Mongo URI	mongodb://localhost:27017/
Mongo Database Name	Crypto
Mongo Collection Name	BTC
SSL Context Service	No value set
Client Auth	REQUIRED
Mode	insert
Upsert	false
Update Query Key	No value set
Update Query	No value set
Update Mode	With whole document
Write Concern	ACKNOWLEDGED

CANCEL

APPLY

When we run these Processors we get the API results in our Consumer and it also gets written down in MongoDB Collection

```
C:\Windows\System32\cmd.exe - .\bin\windows\kafka-console-consumer.bat --topic medium-events --from-beginning --bootstrap-server localhost:...

{"amount": "96728.865", "base": "BTC", "currency": "USD"}
{"data": {"amount": "96724.185", "base": "BTC", "currency": "USD"}}
{"amount": "96724.185", "base": "BTC", "currency": "USD"}
{"data": {"amount": "96724.185", "base": "BTC", "currency": "USD"}}
{"amount": "96724.185", "base": "BTC", "currency": "USD"}
{"data": {"amount": "96724.185", "base": "BTC", "currency": "USD"}}
{"amount": "96724.185", "base": "BTC", "currency": "USD"}
{"data": {"amount": "96711.265", "base": "BTC", "currency": "USD"}}
{"amount": "96711.265", "base": "BTC", "currency": "USD"}
{"data": {"amount": "96735.725", "base": "BTC", "currency": "USD"}}
{"amount": "96735.725", "base": "BTC", "currency": "USD"}
{"data": {"amount": "96735.725", "base": "BTC", "currency": "USD"}}
{"amount": "96735.725", "base": "BTC", "currency": "USD"}
{"data": {"amount": "96735.725", "base": "BTC", "currency": "USD"}}
{"amount": "96735.725", "base": "BTC", "currency": "USD"}
{"data": {"amount": "104230.315", "base": "BTC", "currency": "USD"}}
{"amount": "104806.655", "base": "BTC", "currency": "USD"}
{"data": {"amount": "104230.315", "base": "BTC", "currency": "USD"}}
{"amount": "104806.655", "base": "BTC", "currency": "USD"}
{"data": {"amount": "104819.61", "base": "BTC", "currency": "USD"}}
{"amount": "104819.61", "base": "BTC", "currency": "USD"}
```


POWER BI Dashboard made from Bitcoin Data in MongoDB

