คำอธิบายที่ง่ายๆ สำหรับ Bitcoin คือ Bitcoin เป็นรูปแบบหนึ่งของเงินดิจิทัลหรือเงินเสมือน (Virtual currency) ที่คนสามารถส่งหรือรับเงินผ่านอินเทอร์เน็ตได้ มันคล้ายกับเงินทั่วๆไป แต่มีอยู่เฉพาะในรูปแบบอิเล็กทรอนิกส์เท่านั้น สิ่งที่ทำให้ Bitcoin น่าสนใจคือมันทำงานบนเครือข่ายที่ไม่มีศูนย์กลางหรือที่เรียกว่า Decentralize network ซึ่งหมายความว่าไม่มีหน่วยงานกลาง เช่น รัฐบาลหรือธนาคารที่ควบคุมมันได้

แทนที่จะใช้เหรียญหรือเงินกระดาษ Bitcoin ใช้เทคโนโลยีที่เรียกว่า บล็อกเชน เพื่อบันทึกรายการธุรกรรม บล็อกเชนคือบัญชีสาธารณะที่บันทึกรายการทุกๆ ครั้งอย่างปลอดภัยและโปร่งใส เทคโนโลยีนี้ช่วยให้มีความถูกต้องและปลอดภัยสำหรับสกุลเงิน

คนสามารถซื้อหรือขายสินค้าและบริการด้วย Bitcoin ได้ และมันยังสามารถถูกซื้อขายบนแพลตฟอร์มออนไลน์ต่างๆได้ กระบวนการสร้าง Bitcoin ใหม่เรียกว่า "การทำเหมือง" หรือ Bitcoin Mining ซึ่งเป็นกระบวนการที่คอมพิวเตอร์ที่มีพลังงานสูงแก้ปัญหาคณิตศาสตร์ที่ซับซ้อนเพื่อตรวจสอบและเพิ่มรายการใหม่ลงในบล็อกเชน

Bitcoin ได้รับความสนใจเนื่องจากศักยภาพที่จะให้การเงินที่ไม่มีศูนย์กลางและเสรีภาพทางการเงินมากขึ้น แต่ก็มีความรุนแรงในการเปลี่ยนแปลงราคาและค่าความมั่งคั่งที่มีโอกาสเปลี่ยนแปลงได้มาก หรือมีความ Volatile นั่นเอง

Dataset ที่มีการใช้ในการวิเคราะห์คือ Dataset ที่บันทึก Transaction สำหรับแต่ละ Wallet ID ว่าแต่ละ Wallet ID มี label เป็นอะไร มีกี่ transaction ความสูงของ block เป็นเท่าไร เป็นต้น ซึ่ง dataset จะประกอบไปด้วย 8808 row และ 46 columns

Attribute ที่สำคัญ

Address - คือ wallet ID ที่ unique สำหรับผู้ใช้ 1 คน สามารถูกมองว่าเป็น 1 account ได้

Label – ชนิดของการจัดพวกสำหรับ transaction เช่น อาจจะจัดเป็นนักพนัน gambling หรือนักขุด miner

n\_transaction – จำนวนของ transaction เช่น ใน 1 address สามารถธุรกรรมได้มากกว่า 1 transaction

block height - ความสูงของ block ด้วยความที่ blockchain คือ block ของ trabsaction มาต่อๆกัน ความสูงของหรือ block height จะคือความสูงของ block ก่อนหน้าที่กำลังต่อกันอยู่บน transaction ใดๆสำหรับ address นั้นๆ (block แรกจะเรียกว่า genesis block เมื่อมี transaction ที่ถูกต้อง ได้รับการ confirm แล้วจะมีการต่อกันเป็นทอดๆ จึงทำให้ block height คือความสูงของ block ที่ต่อๆกันมาจนถึง transaction นั้นๆ)

fee – ค่าของการทำธุกรรม เปรียบเทียบได้เหมือนกับที่ทำในธนาคาร

size – ปริมาณของ data ที่ transaction ในแต่ละ address เข้าไป occupied space เช่น bitcoin transaction นี้มีค่าขนาดของ data เท่ากับ 2 byte

weight - คือ block weight เป็น measurement ที่ใช้วัด size อันหนึ่งที่ได้มาจากผลลัพธ์ของ segregated witness (segwit) เป็น idea ที่ upgrade ให้ block ของ bitcoin มี scalability มากขึ้น (รองรับ block ได้มากขึ้น) มองว่ามันคือ software metric อันหนึ่งสำหรับ bitcoin

confirmations - จำนวน block ที่ถูก add เข้าไปใน chain หลังจากที่ transaction นั้นๆได้รับการ verify แล้ว

vin\_sz – จำนวน input ของสำหรับแต่ละ transaction นั้นๆ เอามาจาก output ของ transaction ตัวก่อนมา เช่น ถ้า input เป็น 2 หมายถึง transaction นี้มี 2 input ที่ถูกส่งมาจาก sender ตัวแปร input นี้จะช่วยให้เอามา verify ไปเป็น amount หรือปริมาณเงินที่เอาไปใช้ได้ใน transaction นั้นๆ

vout\_sz - จำนวน output ไอเดียเหมือน input แต่จะออกมาจาก previous transaction เพื่อไปเข้า current transaction แทน

inputsAmount – ไอเดียเหมือน vin\_sz แต่จะ report ออกมาเป็นปริมาณ BTC เลย เช่น vin\_sz มีขนาดเป็น 2 แต่ละหน่วยของมัน represent 1 BTC จะทำให้ vin\_sz มี amount เป็น 2 BTC

outputsAmount – ไอเดียเดียวกัน แต่จะมองเป็นสำหรับของ output แทน (สำหรับ blockchain นั้น data ของ previous block จะออกมาเป็น output เพื่อที่จะเข้าไปป้อนเป็น input ของ current block)