

BAB IV

METODELOGI DAN DATA

IV.1. Metodologi

Metodologi yang digunakan dalam penelitian ini adalah dengan metode statistik deskriptif dan regresi.

IV.1.1. Statistik Deskriptif

Metode ini mengikuti metode riset yang dilakukan oleh Mahabub Hossain (2002). Metode ini mengidentifikasi karakteristik responden anggota LPP UMKM secara umum, dari segi keanggotaan, kondisi infrastruktur lingkungan, serta kondisi sosial dan ekonomi rumah tangga responden. Dalam metode ini juga dilakukan analisa komparatif dengan membandingkan data responden yang disurvei pada saat sebelum menjadi anggota (dari kuesioner Uji Kelayakan pada lampiran 1) dengan data responden yang sama saat disurvei penulis pada bulan April 2008 (dari kuesioner pada lampiran 2). Metode ini menggunakan analisa *mean* (rata-rata) dan standar deviasi.

Mean atau nilai rata-rata atau rerata dapat diartikan sebagai nilai khas yang mewakili suatu himpunan data. Nilai khas cenderung terletak secara terpusat dan merupakan nilai yang mewakili seluruh nilai observasi. Masing-masing rerata digunakan untuk data yang tidak dikelompokkan (data mentah) dan data yang dikelompokkan dalam distribusi frekuensi.

Rerata, dinotasikan dengan \bar{x} , diperoleh dengan menjumlahkan semua data dan membaginya dengan jumlah pengamatan. Sederhananya dapat diilustrasikan sebagai berikut:

$$\bar{x} = \frac{\sum \bar{x}_i}{n}$$
$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad \dots (IV.1)$$

dimana: $i = 1, 2, \dots, n$

Simpangan atau deviasi menunjukkan berapa banyak suatu nilai berbeda dari rata-rata hitungnya. Simpangan dapat bernilai positif apabila nilai yang

dikurangi lebih besar dibandingkan dengan rata-ratanya, dan sebaliknya dapat bernilai negatif apabila nilai yang dikurangi lebih kecil nilainya dibandingkan dengan rata-ratanya. Rata-rata dari deviasi yang dikuadratkan dinamakan **varians**, dan akarnya disebut dengan **standar deviasi**.

- Rumus Varians:

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} \dots (IV.2)$$

- Rumus Standar Deviasi

- a. Pangkat Dua

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}} \dots (IV.3)$$

- b. Rata-rata Pangkat Dua

$$\sigma = \sqrt{\frac{\sum_{i=1}^N X_i^2}{N} - \left(\frac{\sum_{i=1}^N X_i}{N}\right)^2} \dots (IV.4)$$

Untuk sampel, atau yang datanya (N) kurang dari atau sama dengan 30, maka deviasi standar dihitung menggunakan rumus:

$$s = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N - 1}} \dots (IV.5.)$$

IV.1.2. Regresi

Metode regresi digunakan untuk melihat pengaruh beberapa variabel independen terhadap suatu variabel dependen. Pada penelitian ini digunakan variabel dependen yang berupa *dummy*, yaitu miskin atau tidak miskin. Regresi terhadap suatu model dengan variabel *dependent* berbentuk *dummy* disebut dengan *binary probability models*. *Binary probability models* dapat diregresi dengan 3 pendekatan, yaitu: *Linear Probability Model* (LPM), Model Logit, dan Model

Probit. Model yang termudah untuk melakukan pendekatan terhadap *binary probability models* adalah LPM.

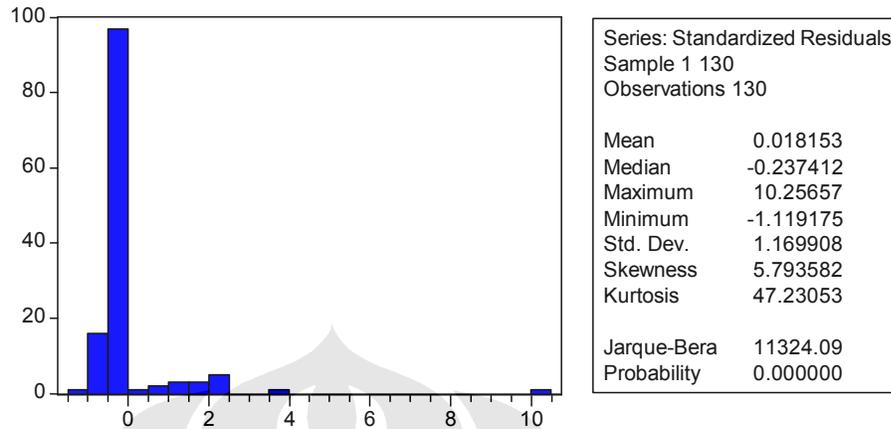
Pada skripsi ini digunakan model logit sebagai acuan analisa ekonomi karena beberapa alasan, diantaranya:

1. Model logit lebih bisa mengestimasi variabel regressor (*independent variabel*) dengan *range* yang lebih besar dibandingkan dengan LPM. *Range* variabel untuk logit bisa mencapai $-\infty < x < \infty$.
2. Pada LPM, kita harus mengasumsikan bahwa errornya terdistribusi normal.
3. Gujarati mengatakan bahwa LPM tidak tepat secara logika. Hal ini dikarenakan probabilita keberhasilan tidak selalu meningkat secara linier terhadap variabel independennya.

Oleh karena itu diperlukan model probabilita yang memiliki beberapa keunggulan dari LPM: (1) ketika X_i meningkat, $P_i = E(Y=1 | X_i)$ meningkat tapi tidak melebihi batas probabilita antara nol dan satu., dan (2) menurut John Aldrich dan Nelson dalam Gujarati (2003), ada hubungan tidak linier antara P_i dan X_i , dimana “*one which approaches zero at slower and slower rates as X_i gets small and approaches one at slower and slower rates as X_i gets very large.*” (John Aldrich dan Nelson dalam Gujarati (2003)).

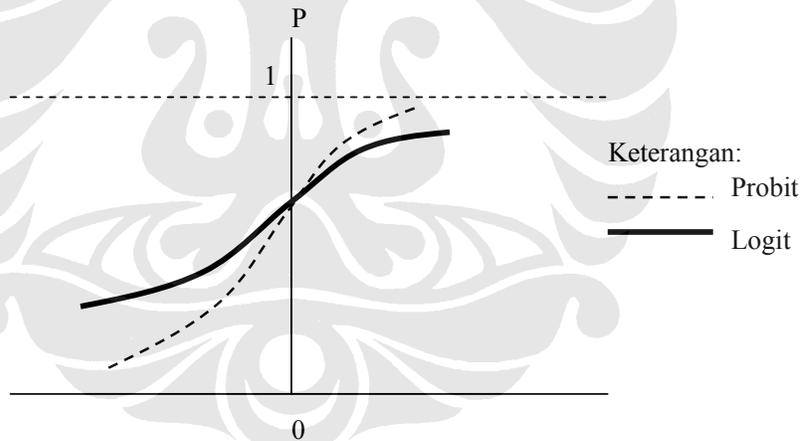
Penelitian ini juga tidak menggunakan model probit karena model ini digunakan ketika diasumsikan errornya (selisih antara nilai variabel yg diestimasi dengan variabel yg sebenarnya) terdistribusi normal. Artinya, nilai error mengikuti distribusi normal dengan mean (nilai rata-rata) bernilai nol dan variansnya konstan bernilai satu ($\text{var}(\epsilon)=1$). Berdasarkan hasil *normality test* yang dilakukan (lihat bagan IV.1.), grafik tidak mendekati bentuk distribusi normal (bentuk lonceng) dan dari *probaility* Jarque-Bera, diyakini bahwa error yang digunakan dalam penelitian ini tidak terdistribusi normal dengan tingkat signifikansi 1%.

Bagan IV. 1. Hasil *Normality Test*



Perbedaan model probit dan logit dapat terlihat dari bagan IV.2., dimana model logit probabilitanya lebih lambat mendekati nol dan satu dibandingkan probit.

Bagan IV. 2. Distribusi Kumulatif Logit dan Probit



Model Logit yang digunakan dalam skripsi ini menggunakan asumsi mean (nilai rata-rata) adalah nol dan variansnya mengikuti distribusi logistik sebesar atau bisa ditulis $\text{var}(\epsilon) = \frac{\pi^2}{3}$ di mana $\pi = 22/7$. Jika nilai koefisien probit (Z) dikalikan dengan $\frac{\pi^2}{3}$ (atau sebesar 1,81) maka akan diperoleh nilai koefisien logit. Model logit lebih dipilih daripada model probit karena secara matematis lebih sederhana (Gujarati, 2003, p. 614).

Logit dengan data individual mirip dengan model regresi *Ordinary Least Squares* (OLS) dengan data silang (*cross section*). Perbedaannya terletak pada variabel dependen dan interpretasi. Pada model logit, variabel dependen terdiri atas bilangan biner 0 dan 1 (mewakili kondisi ya dan tidak). Sedangkan interpretasi atau estimasi pada model logit menunjukkan besarnya kemungkinan suatu kejadian, yang ditunjukkan dengan persentase probabilitas, sehingga nilainya antara 0% hingga 100%.

Persamaan berikut menjelaskan variabel dependen dalam logit:

$$P_i = E(Y = 1|X_i) = \frac{1}{1 + e^{-\beta_1 + \beta_2 X}} \quad \dots(\text{IV.6})$$

Persamaan tersebut dapat disederhanakan menjadi:

$$P_i = \frac{1}{1 + e^{-z}} = \frac{e^z}{1 + e^z} \quad \dots(\text{IV.7})$$

Dengan $z = \beta_1 + \beta_2 X$. Persamaan (IV.6) ini dikenal dengan fungsi distribusi logistik kumulatif (*cumulative logistic distribution function*). z berkisar antara $-\infty$ hingga $+\infty$, sedangkan P_i berkisar antara 0 hingga 1, sehingga P_i tidak linier dengan β dan X . Artinya, persamaan ini tidak bisa diregresi dengan prosedur OLS.

Untuk itu, maka bentuk persamaan dimanipulasi sebagai berikut:

$$1 - P_i = \frac{1}{1 + e^z} \quad \dots(\text{IV.8})$$

Kemudian diubah menjadi:

$$\frac{P_i}{1 - P_i} = \frac{e^z}{1 + e^z} \div \frac{1}{1 + e^z} = \frac{e^z}{1 + e^z} \times \frac{1 + e^z}{1} = e^{z_i} \quad \dots(\text{IV.9})$$

ratio $\frac{P_i}{1 - P_i}$ disebut pula sebagai *odds ratio* (rasio resiko) suatu peristiwa, yaitu rasio kemungkinan terjadinya suatu peristiwa terhadap kemungkinan tidak terjadinya suatu peristiwa.

$$Odds_i = \frac{P_i}{1 - P_i}$$

Nilai odds harus berkisar antara nol hingga positif, tetapi tidak ada batas atas; ketika P_i mendekati 1, maka $P_i/(1 - P_i)$ akan menjadi infinite (tak terhingga). Tetapi ketika akan digunakan sebagai sebuah persamaan linier dengan parameter nilai odds ini masih ada batas bawah bernilai nol, sehingga persamaan tersebut

harus dimanipulasi lagi. Manipulasi tersebut bisa dilakukan dengan mengambil natural logaritma, dengan persamaannya menjadi:

$$\ln Odds_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \ln e^{Z_i} = Z_i$$

Dimana,

$$Z_i = \beta_1 + \beta_2 X \quad \dots(IV.10)$$

Secara lengkap Model Logistic Linier Regression Model (model log odds atau logit) bisa ditulis sebagai berikut ini :

$$\ln \frac{P(Y_i=1)}{1-P(Y_i=1)} = \ln(Odds_i) = \alpha + \sum_{k=1}^K \beta_k X_{ik} = Z_i \quad \dots(IV.11)$$

Untuk menyederhanakan model maka digunakan $Z_i = \alpha + \sum_{k=1}^K \beta_k X_{ik}$. Kemudian kalau kita berikan antilog/diekspensialkan pada kedua sisinya, model tersebut bisa dilihat sebagai nilai odds sukses, yaitu:

$$Odds_i = \frac{P(Y_i=1)}{1-P(Y_i=1)} = \exp(\alpha + \sum_{k=1}^K \beta_k X_{ik}) = \exp(Z_i) = e^{Z_i} \quad \dots(IV.12)$$

Ada beberapa asumsi yg harus dipenuhi dalam menganalisa logit, yaitu:

- 1) Datanya diasumsikan homoskedastis
- 2) Distribusi errornya diasumsikan mengikuti distribusi logistik
- 3) Meannya nol

IV.1.3. Prosedur Pengujian

IV.1.3.1. Uji Signifikansi Variabel secara parsial

Analisis statistik secara parsial digunakan untuk melihat signifikansi dari masing-masing variabel bebas secara individual dalam menjelaskan variabel terikat pada model dengan menggunakan uji z. Pengujian tersebut menggunakan tingkat signifikansi dengan hipotesa sebagai berikut:

H0: variabel independen secara individual tidak mempengaruhi variabel dependen

H1: Variabel independen secara individual mempengaruhi variabel dependen

Ketentuan menerima atau menolak H0 ditentukan melalui probabilita Z hitung (P-Value) masing-masing variabel independen dengan tingkat nyata (α). Jika P value < α maka Ho ditolak dan menerima H1. Tingkat nyata (α) ini dihitung berdasarkan atas nilai tingkat keyakinan. Nilai α dapat dilihat pada tabel berikut:

Tabel IV. 1. Nilai Tingkat Keyakinan dan Tingkat Nyata

Tingkat Keyakinan	Tingkat Nyata
95%	5% = 0,05
90%	10% = 0,10
80%	20% = 0,20

Sumber: Farahnasy (2006) dalam Bagiastomo (2008)

Nilai Odds ratio hasil estimasi menjelaskan bahwa setiap paramater yg diestimasi mengukur perubahan probabilita atau peluang sukses pada variabel dependen karena adanya perubahan variabel independen tertentu sebesar 1 unit (dengan asumsi variabel independen lain konstan).

IV.1.3.2. Uji R²

Seperti yang digunakan dalam OLS, metode Logit juga memiliki uji untuk melihat seberapa besar varians data mampu dijelaskan oleh model. Menurut Gujarati (2003) ukuran *goodness of fit* biasa (R²) bukan menjadi prioritas utama dalam analisa logit. Penggunaan R-square pada Logit harus menggunakan banyak pertimbangan dan kehati-hatian. Menurut UCLA Academic Technology Service (2007) dalam Kharisma (2007) menjelaskan bahwa penggunaan R-square masih diperdebatkan, terbukti dengan dinyatakan R-square untuk *Logit* sebagai *pseudo R-square* (R-square tiruan/palsu).

Oleh karena itu yang paling utama harus diperhatikan adalah nilai koefisien/odds ratio dan signifikansinya lewat LR Test atau Wald Test. Uji R² ini bisa digunakan ketika akan membandingkan dua model berbeda dengan sampel yang sama. Dengan menggunakan *software* E-Views bisa didapatkan nilai *goodness of fit* dengan McFadden R².

Menurut E-Views (1999) dalam Kharisma (2007) dijelaskan bahwa untuk model logit, penggunaan R-square biasa seperti pada OLS tidak lagi relevan, karena itulah nilainya bisa digantikan oleh Mc.Fadden R-square dan Count R-square. Menurut pada penggunaan *software* E-Views, nilai Mcfadden R-Square dan Count R² bisa didapatkan menggunakan rumus berikut ini:

$$\text{Mc.Fadden } R^2 = 1 - \frac{LL_M}{LL_0}$$

dimana LL_m = Log-Likelihood Intercept only

LL_0 = Log-Likelihood Full Model

$$\text{Count } R^2 = \frac{\text{Number of Correct Prediction}}{\text{Total Number of Observation}}$$

IV.1.3.3. Uji Likelihood Ratio Statistik

Uji likelihood ratio statististik (LR Stat) mirip dengan uji F-Test pada OLS biasa. Uji ini bertujuan untuk mengetahui pengaruh variabel independen secara bersama-sama terhadap dependent variabelnya.

“The LR statistic tests the joint null hypothesis that all slope coefficients except the constant are zero and is computed as $-2(\bar{\ell} - \ell)$. This is the analog of the F-statistic in linear regression models and tests the overall significance of the model. The number in parentheses is the degrees of freedom, which is the number of restrictions under test. Note: this test statistic will not be reported if you do not include a constant in the regressor list.

Probability (LR stat) is the p-value of the LR test statistic. Under the null hypothesis, the LR test statistic is asymptotically distributed as a χ^2 variable, with degrees of freedom equal to the number of restrictions under test.” (Eviews (1999) dalam Bagiastomo (2008))

Hipotesa untuk pengujian tersebut adalah sebagai berikut:

H0: Variabel independen secara bersama-sama tidak mempengaruhi variabel dependen (semua parameternya = nol)

H1: Variabel independen secara bersama-sama mempengaruhi variabel dependen (salah satu parameternya \neq nol)

Ketentuan menerima atau menolak H0 ditentukan melalui probabilita LR Stat (P-Value) dengan tingkat nyata (α) tertentu. Jika P value $<$ α maka Ho ditolak dan menerima H1. Nilai α yang biasa digunakan adalah 0.05 atau 0.1

IV.2. Spesifikasi Model

Penulis menggunakan model yang dikembangkan Chowdury, Ghosh dan Wright (2005) di Bangladesh sebagai dasar penelitian karena adanya kemiripan responden di Bangladesh dan Kabupaten Tangerang yang diteliti. Responden di Bangladesh adalah anggota lembaga keuangan mikro yang tinggal di pedesaan di sekitar Dhaka (ibukota Bangladesh), sementara responden yang penulis teliti juga nasabah lembaga keuangan mikro yang mengadopsi sistem Grameen Bank dari Bangladesh dan juga tinggal di pedesaan dekat Jakarta. Yang membedakan adalah kondisi perekonomian masyarakat pedesaan di Bangladesh yang relatif lebih miskin dibanding di Indonesia, khususnya Kabupaten Tangerang. Fasilitas dan infrastruktur di Kabupaten Tangerang juga relatif lebih baik dibanding di wilayah pedesaan Bangladesh, meskipun sama-sama dekat dari Ibukota negara.

Model yang dikembangkan Chowdury, Ghosh dan Wright (2005) digunakan untuk menjelaskan faktor-faktor yang mempengaruhi kemiskinan anggota LPP UMKM, dengan variabel dependen berupa *dummy* (1 = miskin, 0 = tidak). Penelitian ini menggunakan model sebagai berikut:

$$\text{Prob}(P = 1) = f(X_P, X_H, X_V) \quad \dots(\text{IV.1})$$

Dimana P adalah variabel boneka (*dummy*), dengan 1 = miskin dan 0 = tidak. Probabilitas miskin (P=1) penerima kredit mikro dipengaruhi oleh faktor keikutsertaan dalam program kredit mikro (X_P), faktor karakter sosial dan ekonomi individu atau rumah tangga (X_H), dan faktor infrastruktur desa tempat tinggal (X_V). Estimasi yang dipecah-pecah tersebut bertujuan untuk mencegah adanya korelasi antar variabel independen dan konsistensi arah variabel, disamping ingin dilihat pengaruhnya secara parsial tanpa ada gangguan dari faktor lain. X_P diestimasi dengan persamaan berikut:

$$X_P = \beta_0 + \beta_1 T \quad \dots(\text{IV.2})$$

Dimana, T adalah durasi keikutsertaan dalam program kredit mikro LPP UMKM Kabupaten Tangerang (bulan). Model ini untuk melihat dampak lamanya seseorang bergabung dalam program kredit mikro terhadap kemiskinan. Pengaruh durasi keikutsertaan dalam program kredit mikro (LPP UMKM) berdasarkan

Chowdury, Ghosh dan Wright (2005), diharapkan negatif terhadap kemiskinan. Artinya, semakin lama bergabung dengan program kredit mikro, diharapkan semakin kecil probabilitas miskinnya.

X_H diestimasi dengan persamaan berikut:

$$X_H = \beta_0 + \beta_1(E_f, E_m, F_L, M_L) \quad \dots(\text{IV.3})$$

Dimana,

E_f adalah tingkat pendidikan anggota keluarga wanita tertinggi (tahun sekolah)

E_m adalah tingkat pendidikan anggota keluarga pria tertinggi (tahun sekolah)

F_L adalah jumlah anggota keluarga wanita yang termasuk usia angkatan kerja

M_L adalah jumlah anggota keluarga pria yang termasuk usia angkatan kerja.

Model tersebut mengestimasi dampak karakteristik rumah tangga berupa tingkat pendidikan dan anggota keluarga yang berusia angkatan kerja terhadap kemiskinan. Tingkat pendidikan adalah faktor penting dalam pengembangan kegiatan ekonomi suatu negara. Barro dalam Bils & J.Klenopw (2000) yang dikutip oleh Bagiastomo (2008) melalui penelitiannya menemukan fakta bahwa pendidikan dan pertumbuhan ekonomi memiliki hubungan yang positif dan signifikan. Begitu pula dalam skala ekonomi rumah tangga, diharapkan semakin tinggi tingkat pendidikan sebuah keluarga, maka probabilitas miskin semakin kecil. Pengaruh anggota keluarga yang masuk usia angkatan kerja berdasarkan penelitian Chowdury, Ghosh dan Wright (2005) di Bangladesh berpengaruh negatif terhadap kemiskinan. Semakin banyak jumlah anggota keluarga yang termasuk usia angkatan kerja, maka diharapkan tingkat pendapatan keluarga tersebut akan meningkat dan probabilitas miskin semakin kecil.

X_V diestimasi dengan persamaan berikut:

$$X_V = \beta_0 + \beta_1(D_{SD}, D_{SMP}, D_W, D_{El}, JM, JR, JB, JC) \quad \dots(\text{IV.4})$$

Dimana:

D_{SD} = variabel dummy (1 = terdapat SD³⁴ di desanya dan 0 = tidak)

D_{SMP} = variabel dummy (1 = terdapat SMP³⁵ di desanya dan 0 = tidak)

D_W = variabel dummy (1 = terdapat sumber air bersih di desanya dan 0 = tidak)

³⁴ Termasuk Madrasah Ibtidaiyah (MI)

³⁵ Termasuk Madrasah Tsanawiyah (MTs)

D_{El} = variabel dummy (1 = terdapat aliran listrik di desanya dan 0 = tidak)

JM = jarak dari rumah ke pasar terdekat (km)

JR = jarak dari rumah ke jalan beraspal (km)

JC = jarak dari rumah ke Pusat Kabupaten Tangerang (km)

Model X_V mengestimasi pengaruh infrastruktur di lingkungan responden tinggal terhadap kemiskinannya. Faktor infrastruktur didefinisikan dalam keberadaan infrastruktur pendidikan (SD, SMP, SMA), air bersih, listrik PLN, pasar, jalan beraspal, dan perbankan. Schultz dalam Vernon (2002) mengatakan bahwa masyarakat agraris yang tradisional tetap miskin karena sebagian besar negara-negara miskin hanya memberikan kesempatan ekonomi dan teknis yang terbatas, jadilah mereka sebagai "*poor but efficient*". Untuk memberikan kesempatan ekonomi dan teknis yang diperlukan oleh masyarakat miskin tersebut, maka sarana dan prasarana (infrastruktur) mutlak diperlukan. Terdapatnya fasilitas pendidikan, listrik dan air bersih diharapkan dapat memperkecil kemungkinan suatu rumah tangga untuk miskin. Kemudian, semakin dekat rumah dengan pusat perekonomian yang diwakili oleh jalan beraspal, pasar dan pusat Kabupaten Tangerang diharapkan dapat memperkecil kemungkinan rumah tangga tersebut untuk miskin.

IV.3. Data

Data yang digunakan dalam penelitian ini adalah data primer yang penulis survey menggunakan kuesioner (lihat lampiran 2) dan data Uji Kelayakan yang merupakan hasil survey yang dilakukan oleh LPP UMKM sebelum responden menjadi anggota (lihat lampiran 1). Metode model logit Chowdury, Ghosh dan Wright (2005) menggunakan data dari kuesioner yang penulis survey pada bulan April 2008 terhadap 130 responden anggota LPP UMKM. Responden dibagi merata dalam 9 kelompok anggota berdasarkan waktu bergabungnya, dengan pembagian sebagai berikut:

- (i). Kelompok 1 adalah responden yang menjadi anggota antara bulan Juli hingga Desember 2003.
- (ii). Kelompok 2 adalah responden yang menjadi anggota antara bulan Januari hingga Juni 2004.

- (iii). Kelompok 3 adalah responden yang menjadi anggota antara bulan Juli hingga Desember 2004.
- (iv). Kelompok 4 adalah responden yang menjadi anggota antara bulan Januari hingga Juni 2005.
- (v). Kelompok 5 adalah responden yang menjadi anggota antara bulan Juli hingga Desember 2005.
- (vi). Kelompok 6 adalah responden yang menjadi anggota antara bulan Januari hingga Juni 2006.
- (vii). Kelompok 7 adalah responden yang menjadi anggota antara bulan Juli hingga Desember 2006.
- (viii). Kelompok 8 adalah responden yang menjadi anggota antara bulan Januari hingga Juni 2007.
- (ix). Kelompok 9 adalah responden yang menjadi anggota antara bulan Juli hingga Desember 2007.

Tujuan pembagian responden tersebut adalah untuk membandingkan perubahan kesejahteraan anggota berdasarkan lamanya ia bergabung menjadi anggota LPP UMKM.

Dari 9 cabang LPP UMKM, cabang kecamatan Sukadiri adalah yang pertama kali didirikan. Kecamatan ini merupakan satu-satunya kecamatan yang memiliki nasabah yang tergolong dalam kelompok 1 hingga 9 dalam pembagian kelompok di atas. Maka penulis menetapkan cabang kecamatan Sukadiri, Kabupaten Tangerang sebagai area pengambilan sampel. Sampel diambil secara acak dari data Uji Kelayakan yang diambil sebelum anggota bergabung, setelah mengeliminasi sejumlah anggota yang sudah keluar.

Untuk menentukan kemiskinan obyektif responden pada saat sebelum menjadi anggota LPP UMKM (saat Uji Kelayakan), penulis menggunakan garis kemiskinan desa dari data SUSENAS tahun 2003, 2004, 2005 dan 2006. Untuk menentukan pendapatan dan aset dengan harga konstan penulis menggunakan data Indeks Harga Konsumen (IHK) Umum bulanan antara bulan Januari 2003 hingga April 2008 yang dipublikasikan oleh Bank Indonesia.