



Penerapan Regresi Logistik Biner Multilevel pada Partisipasi Angkatan Kerja di Provinsi Jawa Tengah Tahun 2018

Meilani Thereza Saragih^a, Anindia Wahyu Inayah^b, Rani Nooraeni^{c,}, Mikha Aprilio^d, Marita Mutiara Sinsyi^e, Yolanda Rizkie Aprilia^f*

^a Politeknik Statistika STIS, Jalan Otto Iskandardinata No. 64C, Jakarta Timur, 13330.

Email: meilanithereza28@gmail.com

^b Politeknik Statistika STIS, Jalan Otto Iskandardinata No. 64C, Jakarta Timur, 13330.

Email: anndiawahyui@gmail.com

^{c,*} Politeknik Statistika STIS, Jalan Otto Iskandardinata No. 64C, Jakarta Timur, 13330.

Email: raninoor@stis.ac.id

^d Politeknik Statistika STIS, Jalan Otto Iskandardinata No. 64C, Jakarta Timur, 13330.

Email: mikhaaprilio@gmail.com

^e Politeknik Statistika STIS, Jalan Otto Iskandardinata No. 64C, Jakarta Timur, 13330.

Email: maritasinsyi54@gmail.com

^f Politeknik Statistika STIS, Jalan Otto Iskandardinata No. 64C, Jakarta Timur, 13330.

Email: yolandarizkie@gmail.com

ABSTRACT

Logistic regression analysis is divided into multiple types according to its response variable, such as binomial, multinomial, ordinal, and poisson. Logistic regression with binomial response is the simplest one. It is used to analyze the relationship between one dependent binary variable with one or more independent variable in which the response is a dichotomous variable, counting the response as either 1 or 0: counting a 1 if a characteristic is present and counting 0 if a characteristic is not present. However, binary logistic regression is less precise to be used on a case with hierarchical structure. Data with hierarchical structure may have an effect from its parent group. A suitable method to solve a case with hierarchical structure is a multilevel model. This research presents a comparison between single-level binary logistic regression versus multilevel binary logistic regression to analyze 2018's labor force participation rate of Central Java. Variables used in this research are working status as the response variable, type of residential area, gender, age, education, marriage status, and certified training as the first-level explanatory variable, Human Development Index (HDI), Gross Domestic Product (GDP), and minimum wage as second-level explanatory variable. The best model is selected by comparing the deviance on null model at single-level and multilevel with additional check for random effect. The result shows that the null model at multilevel has smaller deviance compared to the null model at one level, and random effect was also observed. Therefore, multilevel binary logistic regression is the best model to analyze 2018's labor force

* Corresponding author.

* Alamat e-mail: raninoor@stis.ac.id

participation rate of Central Java. Parameter estimation with binary logistic regression multilevel shows that type of residential area, gender, age, education, marriage status, and certificated training significantly affected the outcome of 2018's labor force participation rate of Central Java. In addition to that, effect from the differences of district/city stated with Human Development Index (HDI) and Gross Domestic Product (GDP) as contextual variables. The 0,82 percent variation in working status at Central Java Province in 2018 was caused by the characteristic differences between districts/cities of the population.

Keywords: binary logistic regression, multilevel, working status.

ABSTRAK

Analisis regresi logistik dibagi menjadi beberapa jenis data pada responnya, diantaranya yaitu respon biner, multinomial, ordinal dan poisson. Analisis regresi logistik yang paling sederhana digunakan adalah analisis regresi logistik biner. Regresi logistik biner digunakan untuk menganalisa hubungan antara satu variabel respon dan beberapa variabel independen, dengan variabel responnya berupa data kualitatif dikotomi yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik. Namun regresi biner ini kurang tepat apabila digunakan pada kasus yang memiliki struktur hirarki. Data yang memiliki struktur hirarki sebenarnya memiliki efek dari kelompok yang membawahi unit-unit yang diteliti. Suatu metode yang digunakan untuk menyelesaikan kasus berstruktur hirarki adalah model multilevel. Penelitian ini menyajikan perbandingan antara model regresi logistik biner satu level dan model regresi logistik biner multilevel dalam menganalisis partisipasi angkatan kerja di Provinsi Jawa Tengah tahun 2018. Variabel yang digunakan adalah status bekerja sebagai variabel respon, tipe daerah tempat tinggal, umur, pendidikan, status perkawinan, jenis kelamin, dan pelatihan bersertifikat sebagai variabel penjelas level satu serta IPM, PDRB, dan UMK sebagai variabel penjelas level dua. Pemilihan model terbaik dilakukan dengan membandingkan *deviance* pada *null* model regresi logistik biner satu level dan *null* model regresi logistik biner multilevel serta melakukan pengujian signifikan adanya *random effect*. Analisis yang dilakukan menunjukkan hasil bahwa *deviance* pada *null* model regresi logistik biner multilevel lebih kecil dibandingkan *deviance* regresi logistik biner satu level serta terdapat *random effect* sehingga model regresi logistik biner multilevel adalah model yang terbaik dalam menggambarkan partisipasi angkatan kerja di Provinsi Jawa Tengah tahun 2018. Estimasi parameter dengan regresi logistik biner multilevel memperoleh hasil bahwa faktor yang secara signifikan memengaruhi status bekerja di Provinsi Jawa Tengah tahun 2018 adalah tipe daerah tempat tinggal, jenis kelamin, umur, status perkawinan, keikutsertaan dalam pelatihan yang bersertifikat. Selain itu, terdapat pengaruh perbedaan Kabupaten/Kota yang digambarkan dengan IPM dan PDRB sebagai variabel kontekstual. Sebesar 0,82 persen keragaman status bekerja pada Provinsi Jawa Tengah disebabkan oleh perbedaan karakteristik antar kabupaten/kota asal penduduk.

Keywords: multilevel, regresi logistik biner, status bekerja

Diserahkan: 19-03-2020; Diterima: 30-05-2020;

Doi: <https://doi.org/10.29303/emj.v3i1.60>

1. Pendahuluan

Analisis regresi logistik dibagi menjadi beberapa jenis data pada responnya, diantaranya yaitu respon biner, multinomial, ordinal, dan poisson. Analisis regresi logistik yang paling sederhana digunakan adalah analisis regresi logistik biner. Regresi logistik biner digunakan untuk menganalisa hubungan antara satu variabel respon dan beberapa variabel independen, dengan variabel responnya berupa data kualitatif dikotomi yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik (Agresti, 2007). Kini, data dengan respon biner tidak hanya digunakan pada regresi logistik saja, namun sudah mulai berkembang pada metode regresi lainnya.

Namun regresi biner ini kurang tepat apabila digunakan pada kasus yang memiliki struktur hirarki. Struktur data hirarki biasanya berasal dari populasi yang bertingkat atau berjenjang (*hierarchy*) dan berkelompok (*cluster*). Data yang memiliki struktur hirarki adalah data yang terjadi karena individu-individu terkumpul dalam kelompok-kelompok sosialnya. Data yang memiliki struktur hirarki, sebenarnya terdapat efek dari kelompok yang membawahi unit-unit yang diteliti (Hox, 2002). Suatu metode yang digunakan untuk menyelesaikan kasus berstruktur hirarki adalah model multilevel.

Model multilevel merupakan teknik statistik yang telah mengalami pengembangan dari regresi klasik/sederhana. Pengembangan tersebut didasari karena dalam penelitian diberbagai kasus sering dijumpai perbedaan latar belakang pada responden

yang diteliti, sehingga data tersebut memiliki struktur yang bertingkat atau berjenjang (*hierarchy*) dan berkelompok (*cluster*). Goldstein (1995) memperkenalkan pengembangan dari regresi biasa untuk mengatasi permasalahan yang disebabkan dari data yang berstruktur hirarki yaitu analisis *Multilevel Modelling*. Model multilevel merupakan bagian dari *Generalized Linear Mixed Model (GLMM)*, yaitu variabel respon dipengaruhi oleh faktor tetap dan faktor acak. Metode multilevel ini sering digunakan pada kasus sosial, pendidikan dan kesehatan, terutama pada data populasi yang memiliki struktur hirarki. Dalam hal ini yang menjadi fokus penelitian yaitu kasus sosial mengenai ketenagakerjaan.

Informasi ketenagakerjaan merupakan salah satu hal substansial bagi penyusunan kebijakan, monitoring, dan evaluasi dalam rangka pembangunan ekonomi untuk meningkatkan kesejahteraan masyarakat dan pemecahan masalah ketenagakerjaan. Menurut UU No. 13 Tahun 2003 Pasal 1 ayat 1 tentang ketenagakerjaan, “ketenagakerjaan adalah segala hal yang berhubungan dengan tenaga kerja pada waktu sebelum, selama, dan sesudah masa kerja.”

Ketersediaan tenaga kerja sangat dibutuhkan untuk mempercepat terlaksananya rencana pembangunan. Tingkat partisipasi angkatan kerja (TPAK) dapat digunakan untuk mengetahui gambaran penduduk usia kerja yang aktif secara ekonomi. Provinsi Jawa Tengah merupakan provinsi dengan jumlah penduduk bekerja terbesar ketiga setelah Jawa Barat dan Jawa Timur. Namun menurut publikasi BPS Jawa Tengah tahun 2018 masih terdapat kesenjangan antara TPAK laki-laki dan perempuan. Selain itu pendidikan tertinggi yang ditamatkan mayoritas berada pada jenjang pendidikan dasar. Padahal menurut publikasi BPS, IPM Jawa Tengah lebih tinggi dibandingkan IPM Jawa Timur.

Permasalahan utama yang ingin dibahas dalam penelitian ini yaitu membandingkan estimasi parameter menggunakan regresi logistik biner satu level dengan regresi logistik biner multilevel dalam penerapan kasus faktor – faktor yang memengaruhi status pekerjaan di Jawa Tengah. Agar penelitian ini tepat sasaran maka diperlukan pembatas yang jelas sehingga tidak menimbulkan ambiguitas dalam pengambilan data, pengolahan, serta penarikan kesimpulan. Oleh karena itu, penelitian ini hanya dibatasi pada penggunaan regresi logistik biner multilevel. Hal tersebut dikarenakan data tersebut memiliki struktur hirarki, dimana level satu adalah individu di Provinsi Jawa Tengah dan kabupaten/kota di Provinsi Jawa Tengah sebagai level dua.

Melalui penelitian ini diharapkan kedepannya dapat diambil beberapa manfaat diantaranya:

1. Membandingkan antara model estimasi parameter menggunakan metode regresi logistik biner satu level dengan metode regresi logistik biner multilevel dalam penerapan kasus faktor-faktor yang memengaruhi status bekerja di Jawa Tengah.
2. Mengetahui faktor – faktor yang berpengaruh terhadap status bekerja di Provinsi Jawa Tengah berdasarkan model terbaik.
3. Mengetahui pengaruh variasi wilayah terhadap status bekerja di Provinsi Jawa Tengah.

2. Metode Penelitian

Jenis data yang digunakan dalam penelitian ini adalah data sekunder berupa data mentah hasil Survei Angkatan Kerja Nasional (Sakernas) serta data yang sudah diolah dari tabel dinamis yang dikumpulkan oleh Badan Pusat Statistik (BPS) pada 2018. Dari data Sakernas, diambil hasil survei sebanyak 47.508 observasi yang berasal dari provinsi Jawa Tengah dari total sampel 508.460 untuk 34 provinsi Indonesia.

Pada penelitian ini, faktor-faktor yang memengaruhi status bekerja di Provinsi Jawa Tengah dijelaskan secara deskriptif maupun inferensia. Dalam analisis inferensia, digunakan metode analisis regresi logistik biner multilevel (dua level) dan regresi logistik biner satu level yang kemudian dibandingkan dan dipilih model terbaik diantara keduanya

2.1. Analisis Regresi Multilevel

Model regresi multilevel merupakan bagian dari model linear campuran (*Mixed effect model* atau *Generalized Linear Mixed Model (GLMM)*), yaitu suatu model yang menggabungkan komponen tetap atau efek tetap (*fixed effect*) dan komponen acak atau efek acak (*random effect*) ke dalam persamaan model. Model ini digunakan pada data berjenjang (*hierarchy*) yang seringkali ditemukan pada penelitian-penelitian survei di mana unit-unit analisisnya berasalh dari kelompok-kelompok (*cluster*), atau data yang diambil melalui penarikan sampel bertahap (*cluster sampling*).

Penelitian multilevel membagi modelnya menjadi dua, yaitu *null model* dan *conditional model*. *Null model* merupakan kondisi di mana variabel penjelas belum dimasukkan ke dalam model, baik variabel penjelas pada level 1 maupun pada level 2. Sementara itu, *conditional model* merupakan kondisi di mana variabel penjelas, baik penjelas pada model level 1 maupun pada level 2 sudah dimasukkan ke dalam model.

Hox (2010) menyebutkan bahwa model regresi multilevel dapat digolongkan menjadi dua bentuk dasar, yaitu model multilevel dengan *random intercept* dan model multilevel dengan *random slope*. Model multilevel dengan *random intercept* merupakan model di mana *intercept* dimodelkan sebagai *random effect* dari variabel pada level 2 dengan asumsi bahwa setiap kelompok memiliki *intercept* yang berbeda-beda (tidak *fixed* seperti regresi biasa), sedangkan model multilevel dengan *random slope* merupakan model di mana koefisien variabel-variabel penjelas pada level yang lebih rendah dimodelkan sebagai *random effect* dari variabel pada level 2 dengan asumsi bahwa tiap kelompok memiliki kemiringan atau *slope* yang berbeda-beda (tidak *fixed* seperti regresi biasa).

2.2. Analisis Regresi Logistik Biner Multilevel

Dalam model regresi, apabila variabel respon berupa biner atau *dichotomous* biasanya digunakan model regresi logistik yang dalam estimasi parameternya harus menggunakan suatu fungsi penghubung (*link function*). Hal tersebut juga sama diterapkan dalam model multilevel. Apabila variabel respon berdistribusi binomial dengan parameter proporsi (π_{ij}), maka fungsi penghubung yang digunakan adalah logit ($\log\{\pi/(1-\pi)\}$) sehingga modelnya disebut dengan model logistik (Hox, 2002).

Secara umum, model 2 level dengan respon biner dapat dituliskan sebagai berikut (Guo dan Zhao, 2000).

$$\log\left[\frac{\pi_{ij}}{(1-\pi_{ij})}\right] = \beta_0 + \beta_1 X_{1i} + u_j \quad (1)$$

dimana u_j merupakan efek acak pada level 2, tanpa u_j , persamaan (1) menjadi model regresi logistik standar. Bergantung pada u_j , Y_{ij} diasumsikan independen. Seperti dalam kasus pada model linear multilevel, u_j diasumsikan berdistribusi normal, dengan nilai yang diharapkan 0 dan varians σ_u^2 . Model (1) seringkali dideskripsikan sebagai alternatif dalam literatur pada model multilevel dari persamaan berikut.

$$\log\left[\frac{\pi_{ij}}{(1-\pi_{ij})}\right] = \beta_{0j} + \beta_1 X_{1i} + \varepsilon_i \text{ (model level 1)}$$

dan

$$\beta_{0j} = \beta_0 + u_j \quad (3)$$

Secara umum rumusan matematis untuk model *random-intercept* dua level dengan respon biner sebagai berikut.

$$\log\left[\frac{\pi_{ij}}{(1-\pi_{ij})}\right] = \beta_0 + \beta_1 X_{1i} + u_j + \varepsilon_i \quad (4)$$

dimana u_j merupakan *error* pada level 2 yang berdistribusi normal dengan rata-rata 0 dan varians σ_u^2 dan independen terhadap ε_i .

2.3. Tahapan Analisis Regresi Logistik Biner Multilevel

Model regresi logistik biner multilevel yang digunakan adalah model dengan *random intercept* yang mengasumsikan bahwa pengaruh setiap variabel bebas adalah sama pada setiap kabupaten/kota di level dua. *Software* yang digunakan sebagai penunjang dalam pengolahan data pada penelitian ini adalah *software* statistik R dan Stata.

Secara runtut, tahapan analisis regresi logistik biner multilevel dalam penelitian ini adalah sebagai berikut.

a. Pengujian Signifikansi *Random Effect* (*Likelihood Ratio Test*)

Prosedur *maximum likelihood estimator* mampu menghasilkan suatu statistik yang disebut *deviance* yang mampu mengindikasikan seberapa cocok model dengan data (Hox, 2010). Model dengan *deviance* yang lebih rendah akan lebih fit dibandingkan dengan model yang *deviance*-nya lebih tinggi. Pada model bersarang (berhierarki), tes *deviance* dapat digunakan untuk mengetahui apakah model dengan efek *random* lebih baik daripada model tanpa efek *random*. Pengujian dilakukan dengan hipotesis :

$$H_0: \sigma_{u0}^2 = 0 \text{ (efek } random \text{ tidak signifikan)}$$

$$H_1: \sigma_{u0}^2 \neq 0 \text{ (efek } random \text{ signifikan)}$$

Statistik uji :

$$LR = -2\ln\left[\frac{L^{(0)}}{L^{(r)}}\right] \quad (5)$$

dengan $L^{(0)}$ adalah *Likelihood* model logistik tanpa efek *random* dan $L^{(r)}$ adalah *Likelihood* model logistik dengan efek *random*.

Dengan menggunakan $\alpha = 5\%$, maka tolak H_0 jika $LR > \chi_{(a,v)}^2$, dimana v adalah selisih jumlah parameter dari kedua model. H_0 ditolak maka dapat disimpulkan bahwa *random effect* signifikan. Artinya terdapat keragaman atau variasi variabel respons yang signifikan antarkelompok sehingga model multilevel lebih baik dalam menjelaskan data daripada model logistik biasa.

b. Pengujian Signifikansi Parameter secara Simultan

G-test adalah pengujian signifikansi seluruh variabel penjelas di dalam model secara bersama-sama. *G-test* dikenal juga sebagai *likelihood ratio test*, *log-*

likelihood ratio tes, atau Uji G^2 (McDonald, 2014). Hipotesis yang diuji adalah

$$H_0: \gamma_{10} = \gamma_{20} = \dots = \gamma_{p1} = \gamma_{01} = \dots = \gamma_{0q} = 0$$

(tidak ada pengaruh variabel penjelas terhadap variabel respons)

$$H_1: \text{minimal ada } \gamma \neq 0$$

(minimal ada satu variabel penjelas yang berpengaruh terhadap variabel respons)

Statistik uji yang digunakan adalah:

$$LR = -2 \ln \left[\frac{L(\text{null model})}{L(\text{conditional model})} \right] \quad (6)$$

dengan $L(\text{null model})$ adalah Likelihood tanpa variabel penjelas dan $L(\text{conditional model})$ adalah Likelihood dengan variabel penjelas.

G^2 berdistribusi $\chi^2_{(a,r)}$ dimana derajat bebas r adalah jumlah parameter di level 1 dan level 2. H_0 ditolak apabila $G^2 > \chi^2_{(a,r)}$. Ketika H_0 ditolak maka dapat disimpulkan bahwa model dengan variabel penjelas (*conditional model*) fit pada tingkat signifikansi α , atau dengan kata lain pada tingkat kepercayaan $(1-\alpha)$ persen paling tidak terdapat satu variabel penjelas yang memengaruhi variabel respons.

c. Pengujian Signifikansi Parameter secara Parsial

Pengujian parameter dari koefisien regresi secara parsial dapat menggunakan uji Wald (Hosmer et al., 2013). Uji parsial digunakan untuk menentukan variabel bebas mana saja yang memiliki pengaruh terhadap variabel terikat. Hipotesis yang diuji adalah:

$$H_0: \gamma = 0$$

(tidak ada pengaruh variabel penjelas ke- j terhadap variabel respons)

$$H_1: \text{minimal ada } \gamma \neq 0$$

(terdapat pengaruh variabel penjelas ke- j terhadap variabel respons)

Statistik uji yang digunakan adalah:

$$W = \left[\frac{\hat{\gamma}}{Se(\hat{\gamma})} \right] \quad (7)$$

W berdistribusi normal, sehingga H_0 ditolak jika $|W| > Z_{\text{tabel}}$ atau $p\text{-value} < \alpha$. Apabila H_0 ditolak, dapat disimpulkan bahwa variabel penjelas signifikan berpengaruh terhadap variabel respons.

d. Interpretasi Parameter

Interpretasi pada persamaan regresi logistik biner multilevel sama seperti regresi logistik biner satu

level yaitu dapat melalui tanda koefisien regresi, nilai *odds ratio* atau probabilita. Dalam penelitian ini, tanda koefisien regresi dan *odds ratio* digunakan untuk mengetahui arah hubungan dan seberapa besar kecenderungan setiap variabel bebas terhadap pemilihan moda transportasi.

Persamaan *odds ratio* adalah sebagai berikut.

$$\begin{aligned} \widehat{OR} &= \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} \\ &= \frac{\frac{\exp(\widehat{\beta}_0 + \widehat{\beta}_j)}{1 + \exp(\widehat{\beta}_0 + \widehat{\beta}_j)}}{\frac{1}{1 + \exp(\widehat{\beta}_0 + \widehat{\beta}_j)}} \times \frac{1}{\frac{\exp(\widehat{\beta}_0)}{1 + \exp(\widehat{\beta}_0)}} \\ &= \frac{\exp(\widehat{\beta}_0 + \widehat{\beta}_j)}{\exp(\widehat{\beta}_0)} = \exp(\widehat{\beta}_j) \end{aligned} \quad (8)$$

Interpretasi dari \widehat{OR} adalah resiko kecenderungan terjadinya peristiwa $y = 1$ adalah sebesar $\exp(\widehat{\beta}_j)$ kali resiko atau kecenderungan terjadinya peristiwa $y = 1$ pada kategori $x = 0$ (Nachrowi dan Usman, 2002).

e. Interclass Correlation (ICC)

Nilai *interclass correlation* (ICC) menunjukkan tingkat kemiripan karakteristik antar unit mikro (individu) dalam unit makro (kelompok) yang sama (Snijders & Bosker, 1999). Nilai ICC dapat diperoleh dari hasil estimasi variansi pada null model / model yang tidak mengandung variabel bebas. Nilai ICC dapat dihitung melalui:

$$ICC = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2} \quad (9)$$

dengan σ_0^2 adalah variansi antar kelompok pada level dua dan σ^2 adalah variansi antar individu pada level satu.

2.4. Status Bekerja

Survei Sakernas dilakukan pada sekitar 200.000 rumah tangga yang mencakup semua individu dalam rumah tangga, sehingga individu dalam rumah tangga ditetapkan sebagai level 1 pada hierarki multilevel dengan variabel Status Bekerja (bekerja diberi kode 1 dan tidak bekerja diberi kode 0) sebagai variabel dependen (Y). Dalam menentukan status bekerja, penelitian ini menggunakan konsep menurut Badan Pusat Statistik (BPS), yaitu **bekerja** adalah kegiatan ekonomi yang dilakukan oleh seseorang dengan maksud memperoleh atau membantu memperoleh pendapatan atau keuntungan, paling sedikit 1 jam (tidak terputus) dalam seminggu yang lalu. Kegiatan tersebut termasuk pola kegiatan pekerja tak dibayar yang membantu dalam suatu usaha/kegiatan ekonomi.

Penelitian oleh Fanny Ayu Octaviana dengan judul *Pemodelan Status Bekerja Ibu Rumah Tangga Menggunakan Model Multilevel Dengan Respon Biner* yang bertujuan untuk mengetahui hubungan antara status bekerja ibu rumah tangga berdasarkan variabel-variabel pada setiap level dan model yang sesuai dari penelitian tentang status bekerja ibu rumah tangga. Variabel respons yang digunakan ialah Status Bekerja Ibu Rumah Tangga (Y), sedangkan variabel penjelas terdiri dari Pendidikan Terakhir, Usia, dan Pengeluaran Rumah Tangga (level 1), serta Gaji Tenaga Kerja Wanita (level 2). Dari hasil analisis dapat diperoleh kesimpulan bahwa model yang terbaik untuk kasus status bekerja ibu rumah tangga di Indonesia adalah menggunakan model regresi 2-level dengan respon biner yang mengikutsertakan variabel rata-rata gaji tenaga wanita karena nilai Deviance yang dihasilkan paling kecil dibandingkan dengan model lainnya. Artinya, keputusan ibu rumah tangga untuk ikut serta dalam dunia kerja dipengaruhi oleh variabel pendidikan terakhir dan variabel pengeluaran rumah tangga (level 1/individu), serta dipengaruhi oleh adanya daerah tempat tinggal (level 2/kecamatan).

Kemudian, penelitian oleh Rizky Amalia Yulianti dan Vita Ratnasari yang berjudul *Pemetaan dan Pemodelan Tingkat Partisipasi Angkatan Kerja (TPAK) Perempuan di Provinsi Jawa Timur dengan Pendekatan Model Probit*. Dalam penelitiannya, variabel respons yang digunakan adalah hasil pengelompokan dari analisis *cluster* antara TPAK perempuan semua sektor. Sedangkan, variabel penjelas yang digunakan meliputi Tingkat Pendidikan Tamat SD, Tingkat Pendidikan Tamat SMP, Tingkat Pendidikan Tamat SMA, Usia, Persentase Perempuan Berstatus Menikah, Pengeluaran Perkapita, Upah Minimum Kabupaten/Kota, Indeks Pembangunan Manusia (IPM), Daerah Tempat Tinggal Asal Kota, dan Produk Domestik Regional Bruto (PDRB). Berdasarkan hasil analisis kelompok terdapat empat kelompok bentukan yaitu kelompok wilayah TPAK perempuan di sektor pertanian, perdagangan, industri perdagangan, dan pertanian perdagangan. Hasil pemodelan probit dengan menggunakan sepuluh variabel yang diduga memengaruhi TPAK perempuan adalah variabel pengeluaran perkapita sebulan, tenaga kerja perempuan asal kota, dan PDRB yang berpengaruh secara signifikan terhadap TPAK perempuan.

Secara kontekstual, status bekerja tidak hanya dipengaruhi oleh faktor individual, tetapi juga dapat dipengaruhi oleh konteks masyarakat atau kelompok di mana individu tersebut menetap. Faktor kontekstual yang digunakan dalam penelitian ini meliputi Produk Domestik Regional Bruto (PDRB) Perkapita, Indeks Pembangunan Manusia (IPM) di

tiap kabupaten/kota, dan Upah Minimum Kabupaten/Kota. Penggunaan PDRB perkapita sebagai faktor kontekstual karena variabel ini bisa menggambarkan kuantitas dan kondisi perekonomian masyarakat di suatu daerah, sedangkan IPM menjelaskan bagaimana penduduk pada suatu daerah dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan, dan sebagainya. Kemudian, upah minimum regional merupakan suatu standar minimum yang digunakan oleh para pengusaha pada suatu daerah untuk memberikan upah kepada pegawai, karyawan, atau buruh di dalam lingkungan usaha atau kerjanya yang nantinya dapat memengaruhi kesempatan kerja dan pengangguran.

Secara ringkas, variabel penelitian yang digunakan dapat dilihat pada Tabel 1 di bawah ini.

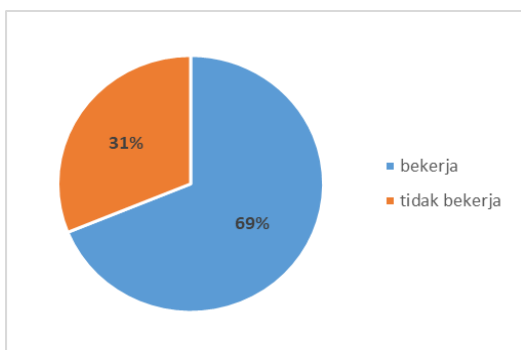
Tabel 1 – Variabel Penelitian

Variabel	Indikator	Keterangan	Skala
Y	Status Bekerja	0 : Tidak Bekerja 1 : Bekerja	Nominal
<i>Level 1 (Individu)</i>			
X_1	Tipe Daerah	0 : Desa 1 : Kota	Nominal
X_2	Jenis Kelamin	0 : Laki - Laki 1 : Perempuan	Nominal
X_3	Status Perkawinan	0 : Belum Kawin 1 : Pernah Kawin	Nominal
X_4	Pernah Mengikuti Pelatihan Bersertifikat	0 : Tidak 1 : Ya	Nominal
X_5	Pendidikan	0 : Di bawah SMP 1 : SMP Ke Atas	Nominal
X_6	Umur	-	Rasio
<i>Level 2 (Kabupaten/ Kota)</i>			
Z_1	Produk Domestik Regional Bruto harga berlaku	-	Rasio
Z_2	Indeks Pembangunan Manusia	-	Rasio
Z_3	Upah Minimum Kabupaten/ Kota	-	Rasio

3. Analisis dan Pembahasan

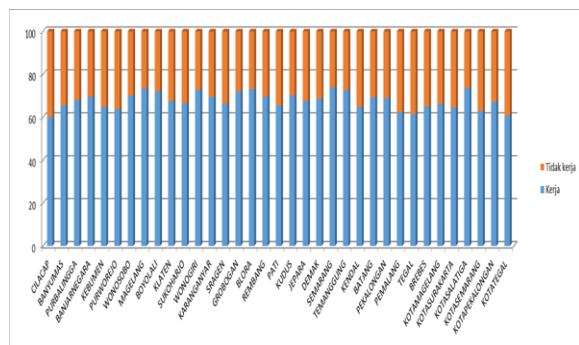
3.1. Gambaran Umum Status Bekerja Penduduk Provinsi Jawa Tengah

Hasil pengolahan data Sakernas Agustus 2018, diketahui bahwa dari total 47.508 sampel penduduk usia kerja di Jawa Tengah, terdapat 69% penduduk yang bekerja dan terdapat 31% penduduk yang tidak bekerja.



Gambar. 1 – Persentase Status Bekerja di Provinsi Jawa Tengah Tahun 2018

Jika ditinjau per kabupaten/kota yang ada di Provinsi Jawa Tengah, dapat diketahui bahwa proporsi penduduk yang bekerja di seluruh kabupaten/kota lebih tinggi dibandingkan proporsi penduduk yang tidak bekerja. Proporsi penduduk yang bekerja di setiap kabupaten/kota di Provinsi Jawa Tengah berkisar antara 60 sampai 75 persen.



Gambar. 2 – Persentase Penduduk yang Bekerja menurut Kabupaten/ Kota di Jawa Tengah

Gambar 2 menunjukkan bahwa Kabupaten Semarang (73,90%), Kota Salatiga (73,43%), Kabupaten Magelang (73,34%), Kabupaten Blora (73,12%), dan Kabupaten Wonogiri (72,70%) adalah lima kabupaten/kota dengan persentase penduduk bekerja tertinggi di Jawa Tengah. Sedangkan persentase terendah terdapat di Kota Semarang (62,83%), Kabupaten Pemalang (62,26%), Kabupaten Tegal (61,47), Kota Tegal (60,75%), dan

Kabupaten Cilacap (60,15%). Perbedaan cakupan antar kabupaten/kota ini disebabkan oleh perbedaan karakteristik wilayahnya, baik dari segi IPM, PDRB, dan UMK.

Berdasarkan faktor sosial-demografi nya, kebanyakan penduduk di provinsi Jawa Tengah bertempat tinggal di perkotaan (55 persen), berjenis kelamin perempuan (52 persen), berstatus pernah kawin (80 persen), berpendidikan tamat di atas SMP (53 persen), dan pernah mengikuti pelatihan bersertifikat (92 persen). Hasil karakteristik penduduk di Provinsi Jawa tengah secara lengkap disajikan pada Tabel 1 sebagai berikut.

Tabel 2 – Profil Penduduk di Provinsi Jawa Tengah tahun 2018 menurut Variabel Penjelas Individual dan Kontekstual

Karakteristik	Jumlah Sampel	Persentase
Daerah Tempat Tinggal		
Perkotaan	26018	55,0
Perdesaan	21040	45,0
Jenis Kelamin		
Laki-laki	22773	48,0
Perempuan	24285	52,0
Status Perkawinan		
Pernah Kawin	37804	80,0
Belum Kawin	9254	20,0
Pendidikan		
SMP ke Bawah	22272	47,0
Di Atas SMP	24786	53,0
Mengikuti Pelatihan		
Ya	3851	92,0
Tidak	43207	8,0

Sumber: raw data Sakernas 2018 (diolah)

3.2. Pemilihan Model Terbaik

Sebelum membentuk model, dilakukan pemilihan dengan memeriksa *random effect* untuk mengkaji apakah terdapat perbedaan grup/kelompok pada kabupaten/kota di Jawa Tengah. Pemeriksaan *random effect* dapat dilakukan dengan cara membandingkan nilai *deviance* antara model regresi logistik biner satu level dengan model regresi logistik biner multilevel pada kondisi tidak mengandung variabel bebas (*null model*) atau dengan melakukan pengujian signifikansi *random effect*. Hasil perbandingan nilai *deviance* adalah sebagai berikut.

Tabel 3 – Hasil Estimasi Nilai Deviance pada Null Model

Model	Deviance
Regresi logistik biner satu level	59279,0
Regresi logistik biner multilevel	59074,0

Tabel 3 menunjukkan bahwa pada kondisi *null model*, nilai *Deviance* pada regresi logistik biner multilevel lebih kecil daripada dengan nilai *Deviance* regresi logistik biner satu level. Selain itu, berdasarkan pengujian signifikansi *random effect*, diperoleh nilai *p-value* sebesar 0,00 dengan $\alpha = 5\%$ dan nilai *Likelihood Ratio Test* sebesar 186,90. Nilai *p-value* yang diperoleh lebih kecil dari α dan nilai *Likelihood Ratio Test* lebih besar dari $\chi^2_{0,05;1} = 3,84$ sehingga dapat disimpulkan dengan tingkat kepercayaan 95% terdapat *random effect* yang signifikan. Dengan demikian, dapat diartikan bahwa terdapat variasi status bekerja antar kabupaten/kota dan model regresi logistik biner multilevel yang mengandung *random effect* lebih baik untuk digunakan dalam menjelaskan data dibandingkan dengan model regresi logistik biner satu level.

3.3. Faktor-Faktor Yang Memengaruhi Status Bekerja di Provinsi Jawa Tengah

Uji Simultan

Pengujian signifikansi parameter secara serentak dilakukan untuk mengetahui apakah model regresi logistik biner multilevel dengan melibatkan variabel penjelas (*conditional model*) lebih baik daripada model tanpa melibatkan variabel penjelas (*null model*). Pengujian signifikansi parameter diperoleh dengan,

$$\begin{aligned} G &= -2 \times \left[\frac{\log \text{likelihood null model}}{\log \text{likelihood conditional model}} \right] \\ &= -2 \times [-29546,085 - (26000,351)] \\ &= 7091,468 \end{aligned}$$

Karena nilai G lebih besar dari $\chi^2_{0,05;9} = 16,919$ sehingga diperoleh keputusan tolak H_0 dan dapat disimpulkan bahwa dengan tingkat kepercayaan 95% terdapat minimal satu variabel penjelas yang memengaruhi status bekerja di Provinsi Jawa Tengah.

Uji Parsial

Pengujian signifikansi parameter secara parsial dilakukan untuk mengetahui pengaruh masing-masing variabel penjelas terhadap status bekerja di Provinsi Jawa Tengah, baik variabel penjelas level 1

maupun level 2. Berikut adalah hasil estimasi parameter dalam model regresi logistik biner multilevel dengan *random intercept*.

Tabel 4 – Hasil Estimasi Parameter Secara Parsial, Signifikansi, dan Nilai Odds Ratio

Variabel	Koefisien	Standard Error	p-value	Odds Ratio
Konstanta	-0,608	0,535	0,256	0,5445
Tipe Daerah Perkotaan Pedesaan	-0,071	0,0235	0,003*	0,9319
Jenis Kelamin Perempuan Laki-laki	-1,389	0,0234	0,000*	0,2493
Umur	-0,021	0,0009	0,000*	0,9795
Status Perkawinan Pernah Kawin Belum Kawin	2,051	0,0350	0,000*	7,7777
Pendidikan \geq SMP $<$ SMP	-0,182	0,0265	0,000*	0,8340
Mengikuti Pelatihan Bersertifikat Pernah Tidak Pernah	0,533	0,0438	0,000*	1,7048
IPM	0,018	0,0074	0,015*	1,0182
PDRB	-0,036	0,0152	0,019*	0,9650
UMK	$2,04 \times 10^{-7}$	$2,04 \times 10^{-7}$	0,395	1

Keterangan: * adalah signifikan pada $\alpha = 5\%$

Penelitian ini menggunakan model regresi logistik biner multilevel dengan *random intercept* karena mengasumsikan pengaruh setiap variabel penjelas terhadap variabel respon sama untuk setiap kabupaten/kota sehingga estimasi parameter *conditional model* menghasilkan persamaan regresi logistik biner multilevel sebagai berikut.

$$\begin{aligned} \ln \left(\frac{\pi_{ij}}{1-\pi_{ij}} \right) &= -0,608 - 0,071 \text{ tipe daerah}_{ij} - \\ &1,389 \text{ jenis kelamin}_{ij} - 0,021 \text{ umur}_{ij} + \\ &2,051 \text{ status perkawinan}_{ij} - \\ &0,182 \text{ pendidikan}_{ij} + \\ &0,533 \text{ pelatihan}_{ij} + 0,018 \text{ IPM}_{ij} - \\ &0,036 \text{ PDRB}_{ij} + 2,04 \times 10^{-7} \text{ UMK}_{ij} \end{aligned}$$

Hasil pengujian signifikansi parameter secara parsial menunjukkan bahwa secara individual, variabel sosio-demografi seperti tipe daerah tempat tinggal, jenis kelamin, umur, dan status perkawinan memengaruhi status bekerja penduduk di Provinsi Jawa Tengah. Selain itu keikutsertaan dalam pelatihan yang bersertifikat juga berpengaruh secara signifikan.

Tipe daerah berpengaruh signifikan terhadap status bekerja dengan nilai koefisien yaitu $-0,071$. Rasio kecenderungan yang diperoleh sebesar $0,9319$ artinya penduduk yang tinggal di daerah perkotaan memiliki kecenderungan untuk bekerja sebesar $0,9319$ kali dibanding penduduk yang tinggal di daerah pedesaan.

Jenis kelamin memberikan berpengaruh signifikan terhadap status bekerja. Rasio kecenderungan perempuan untuk bekerja sebesar $0,2493$ dibandingkan laki-laki. Atau dapat dinyatakan bahwa laki-laki memiliki kecenderungan $4,011$ kali lebih besar dibandingkan perempuan untuk bekerja.

Selanjutnya, umur memiliki pengaruh yang signifikan dan negatif terhadap status bekerja penduduk Jawa Tengah dengan nilai koefisien sebesar $-0,021$. Hal ini berarti bahwa ketika umur seseorang bertambah satu tahun, maka kecenderungan untuk bekerja menjadi $0,9795$ kali dengan asumsi variabel lain konstan. Dengan kata lain, semakin tua seseorang, maka kecenderungan untuk bekerja akan semakin kecil.

Ditinjau dari sisi status perkawinan, penduduk di Provinsi Jawa Tengah yang pernah kawin, dalam hal ini yang dimaksud adalah penduduk yang berstatus kawin, cerai hidup dan cerai mati memiliki kecenderungan sebesar $7,7777$ kali untuk bekerja dibanding penduduk yang belum kawin.

Pendidikan berpengaruh secara signifikan terhadap status bekerja. Nilai koefisien sebesar $-1,182$ dengan rasio kecenderungan sebesar $0,8340$ menunjukkan bahwa penduduk yang pendidikannya di atas SMP memiliki kecenderungan $0,8430$ kali untuk bekerja dibandingkan yang pendidikannya SMP kebawah.

Pelatihan juga berpengaruh secara signifikan dan positif terhadap status bekerja. Penduduk yang pernah mengikuti pelatihan bersertifikat mempunyai kecenderungan $1,7048$ kali untuk bekerja dibandingkan penduduk yang tidak pernah mengikuti pelatihan.

Dari segi kontekstual, penelitian ini menemukan bahwa IPM dan PDRB berpengaruh secara signifikan terhadap status bekerja. Dalam keadaan variabel bebas lainnya konstan, setiap peningkatan 1 poin capaian IPM kabupaten/kota akan meningkatkan kecenderungan penduduk kabupaten/kota tersebut untuk bekerja sebesar $1,0182$. Berkebalikan dengan

IPM, koefisien PDRB yang bernilai negatif menunjukkan bahwa semakin tinggi PDRB, maka kecenderungan untuk bekerja akan semakin berkurang. Hasil penelitian ini juga menemukan bahwa Upah Minimum Kabupaten/Kota (UMK) tidak berpengaruh signifikan terhadap status bekerja penduduk di Provinsi Jawa Tengah.

3.4. Pengaruh Variasi Wilayah Terhadap Status Bekerja

Interclass Correlation (ICC) digunakan untuk menggambarkan besarnya variasi status bekerja antar kabupaten/kota di Provinsi Jawa Tengah. Hasil ICC yang diperoleh sebesar $0,0082016$ artinya sebesar $0,82$ persen keragaman status bekerja pada Provinsi Jawa Tengah disebabkan oleh adanya perbedaan karakteristik antar kabupaten/kota.

4. Kesimpulan dan Saran

4.1. Kesimpulan

Penelitian ini menunjukkan bahwa model yang dihasilkan oleh model regresi logistik biner multilevel lebih baik dalam menggambarkan status bekerja penduduk di Provinsi Jawa Tengah tahun 2018 dibandingkan dengan regresi logistik biner satu level. Hal ini ditunjukkan melalui nilai *deviance* pada regresi logistik biner multilevel lebih kecil daripada nilai *deviance* pada regresi logistik biner biasa.

Penerapan Regresi Logistik Biner Multilevel menyatakan bahwa status bekerja penduduk di Provinsi Jawa Tengah secara signifikan dipengaruhi oleh tipe daerah tempat tinggal, jenis kelamin, umur, status perkawinan, keikutsertaan dalam pelatihan yang bersertifikat. Selain itu, terdapat pengaruh perbedaan Kabupaten/Kota yang digambarkan dengan IPM dan PDRB sebagai variabel kontekstual. Sebesar $0,82$ persen keragaman status bekerja pada Provinsi Jawa Tengah disebabkan oleh perbedaan karakteristik antar kabupaten/kota asal penduduk

4.2. Saran

Berdasarkan hasil analisis dan kesimpulan, saran yang dapat diberikan untuk pemerintah Provinsi Jawa Tengah agar menciptakan lapangan pekerjaan yang semakin banyak. Selain itu, pemerintah dapat membuat kebijakan-kebijakan yang dapat mempermudah penduduk Jawa Tengah untuk memperoleh pendidikan yang lebih tinggi. Misalnya memberikan kesempatan para pekerja untuk menjadi tugas belajar atau kejar paket. Dengan pendidikan yang lebih tinggi, maka kualitas tenaga kerjanya juga akan semakin baik.

Kemudian apabila dilihat dari sisi metodologi, pemodelan status bekerja penduduk Provinsi Jawa Tengah dalam penelitian ini dibatasi dengan menggunakan model *random intercept*. Untuk penelitian selanjutnya, dapat mempertimbangkan untuk melakukan pemodelan menggunakan model regresi logistik biner multilevel dengan *random slope* jika diasumsikan terdapat perbedaan pengaruh variabel penjelas terhadap variabel respon untuk setiap kelompok.

Penelitian ini juga hanya membagi wilayah hingga tingkat kabupaten/kota, sehingga untuk penelitian lebih lanjut diharapkan dapat melakukan pemodelan regresi logistik tiga level, dimana level 1 pada individu, level 2 tingkat kecamatan, dan level 3 tingkat kabupaten/kota.

DAFTAR PUSTAKA

- Adwiluvito, H. (2019). Determinan Pemilihan Moda Transportasi Pekerja Komuter Jabodetabek dengan Model Regresi Logistik Multinomial Multilevel. *Indonesian Journal of Statistics and Its Applications*, 3(1), 49-61.
- Aliyudin, F. & Budyanra. (2016). Faktor-Faktor yang Memengaruhi Komplikasi Persalinan Wanita Usia Subur di Indonesia Menggunakan Data SDKI 2012 (Aplikasi Analisis Regresi Logistik Biner Multilevel). *Jurnal Aplikasi Statistika & Komputasi Statistik*, 8(2), 33-46.
- Badan Pusat Statistik. (2018). *Indeks Pembangunan Manusia 2018*. Jakarta: BPS.
- Badan Pusat Statistik. (2019). *Keadaan Angkatan Kerja Provinsi Jawa Tengah 2018*. Semarang: BPS.
- Badan Pusat Statistik. (n.d.). *Indeks Pembangunan Manusia*. Retrieved Februari 18, 2020 from <https://www.bps.go.id/subject/26/indeks-pembangunan-manusia.html>
- Harlan, J. (2016). *Analisis Multilevel*. Depok: Gunadarma.
- Octaviana, F. (2017). *Pemodelan Status Bekerja Ibu Rumah Tangga Menggunakan Model Multilevel Dengan Respon Biner* [Master's thesis, Institut Teknologi Sepuluh Nopember]. ITS Repository. <http://repository.its.ac.id/id/eprint/2093>
- Panorama, M. & Lemiyana. (2017). Pengaruh Upah Minimum Kota (UMK) Terhadap Kesempatan Kerja dan Pengangguran di Kota Palembang Tahun 2004-2014. *I-Finance*, 3(2), 141-160.
- Yulianti, R.A. & Ratnasari, V. (2013). Pemetaan dan Pemodelan Tingkat Partisipasi Angkatan Kerja (TPAK) Perempuan di Provinsi Jawa Timur dengan Pendekatan Model Probit. *Jurnal Sains dan Seni POMITS*, 2(2), 159-164.